

Tue Jan 20 17:43:17 2004

US-10-022-025a-1.rn1

Page 4

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 09:48:47 ; Search time 124 Seconds
(without alignments)
7151.119 Million cell updates/sec

Title: US-10-022-025a-1
Perfect score: 2009
Sequence: 1 aataataataataatgta.....atagagatgagatatt 2009

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgm2_6/prodata/1/ina/SA_COMB.seq:*
2: /cgm2_6/prodata/1/ina/SB_COMB.seq:*
3: /cgm2_6/prodata/1/ina/SA_COMB.seq:*
4: /cgm2_6/prodata/1/ina/SB_COMB.seq:*
5: /cgm2_6/prodata/1/ina/PCUS_COMB.seq:*
6: /cgm2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	420.8	20.9	1611	3	US-08-948-564-7 Sequence 7, Appli
2	200.4	10.0	1799	2	US-08-560-998-3 Sequence 3, Appli
3	123.2	6.1	1704	3	US-08-948-564-17 Sequence 17, Appli
4	101.2	5.0	1708	3	US-08-991-677-1 Sequence 1, Appli
5	98.2	4.9	275	4	US-09-313-294A-2591 Sequence 2591, Ap
6	89	4.4	1788	3	US-08-948-564-9 Sequence 9, Appli
7	88.8	4.4	1893	1	US-08-532-065B-1 Sequence 1, Appli
8	84.6	4.2	2174	3	US-08-606-505B-63 Sequence 63, Appli
9	84.6	4.2	2174	3	US-08-606-505B-63 Sequence 63, Appli
10	83.2	4.1	1761	3	US-08-033-055A-3 Sequence 3, Appli
11	83.2	4.1	1927	3	US-08-606-505B-64 Sequence 64, Appli
12	83.2	4.1	1927	3	US-08-606-505B-64 Sequence 64, Appli
13	81.4	4.1	5156	2	US-09-091-432-3 Sequence 3, Appli
14	81.4	4.1	5156	2	US-09-387-663-3 Sequence 3, Appli
15	79.2	3.9	1722	3	US-09-033-055A-1 Sequence 1, Appli
16	77.2	3.8	1755	1	US-07-912-900-29 Sequence 29, Appli
17	77.2	3.8	1755	1	US-08-285-309-29 Sequence 29, Appli
18	77.2	3.8	1755	2	US-08-502-046-29 Sequence 29, Appli
19	76.4	3.8	1812	1	US-08-313-075A-37 Sequence 37, Appli
20	76.4	3.8	1824	3	US-08-606-505B-1 Sequence 1, Appli
21	76.4	3.8	1824	3	US-08-606-505B-1 Sequence 1, Appli
22	75.6	3.8	1834	4	US-09-615-192A-232 Sequence 232, App
23	75	3.7	1835	4	US-09-564-808-5 Sequence 5, Appli
24	75	3.7	1838	2	US-09-091-432-1 Sequence 1, Appli
25	75	3.7	1838	4	US-09-387-663-1 Sequence 1, Appli
26	73	3.6	1620	4	US-09-627-216A-11 Sequence 11, Appli

28	72.8	3.6	543	4	US-09-615-192A-230 Sequence 230, App
29	71.8	3.6	1884	4	US-09-564-808-3 Sequence 3, Appli
30	70.2	3.5	1657	3	US-08-948-564-11 Sequence 11, Appli
31	69.2	3.4	1806	4	US-09-351-229-3 Sequence 3, Appli
32	68.6	3.4	1880	4	US-09-564-808-1 Sequence 1, Appli
33	66.6	3.3	1781	4	US-09-499-302A-1 Sequence 1, Appli
34	65.4	3.3	1812	1	US-07-912-900-28 Sequence 28, Appli
35	65.4	3.3	1812	1	US-08-285-309-28 Sequence 28, Appli
36	65.4	3.3	1812	2	US-08-502-046-28 Sequence 28, Appli
37	64	3.2	1506	3	US-09-158-767-7 Sequence 7, Appli
38	64	3.2	1506	3	US-09-158-767-8 Sequence 8, Appli
39	64	3.2	1506	3	US-09-158-767-9 Sequence 9, Appli
40	64	3.2	2261	3	US-09-126-420A-1 Sequence 1, Appli
41	63	3.1	1737	4	US-08-948-564-13 Sequence 13, Appli
42	63	3.1	1824	3	US-08-313-075A-55 Sequence 55, Appli
43	62	3.1	1634	4	US-09-126-420A-2 Sequence 2, Appli
44	61.4	3.1	1644	3	US-08-948-564-5 Sequence 5, Appli
45	61.4	3.1	1644	3	US-08-948-564-5 Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-948-564-7
Sequence 7, Application US/08948564
Patent No. 6121512
GENERAL INFORMATION:
APPLICANT: Siminsky, Balazs
APPLICANT: Tewey, Ralph E.
APPLICANT: Corlin, Frederick T.
TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512ch Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,564
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCE/DOCKET NUMBER: 5051-409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-854-1400
TELEFAX: 919-854-1401
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1611 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 20..1588
US-08-948-564-7
Query Match 20.9% Score 420.8; DB 3; Length 1611;
Best Local Similarity 58.4%; Pred. No. 4.9e-107;
Matches 1029; Conservative 0; Mismatches 487; Indels 256; Gaps 6;

Tue Jan 20 17:43:17 2004

us-10-022-025a-1.rnt

Page 2

QY 221 AAGCAACCAACCTGCTCTCTCCCTCCGCGCCGCAATCATCTGCTGCGCATATCT 280
DB 74 ACCCAAGAAACCTTGCAATGAGTCTCTTGTATCATGAGCTCATCTGCTTAACTACT 133
QY 281 CTCTCTTATGACCTATCCGCTGAGCTCTGCTGAGGAAATACCTCTGCGCGCTTA 340
DB 134 TCTATTTACCTGCTACACCCCGGTGCTCTGCTGAGGCAAGTACTACACC-----TA 186
QY 341 AATCGGCTTCAATACAAACCGGAAACGTTATTCGCGCTCCAAAGGCTTCCCTTGGT 400
DB 187 CTCTCCCCCTTTCAA-----TCATTCGCGCTCCAAAGGCTTCTCTTAAT 235
QY 401 GGAAGCATCTCATCTCATGCAAGCACTTACGCTACCGAATCGCTGATGACGCTAG 460
DB 236 GGAAGCATGAGGCTCATGACTTCC---CTGCGCCATCACTGATGACGCGCGCGCC 292
QY 461 AATTTGGAAGCAAGAGCTCATGCTTCACTGCTTACGCTGAGAGAGCTCGCTGATGCTACG 520
DB 293 AATGCAAGAGCAAGAGCTCATGCTTCACTGCTTACGCTGAGAGAGCTCGCTGATGCTACG 352
QY 521 TGGCAATCCGCACTGACGAAAGATCTGAAATGACCCGCTTTTGTGATGACGCTG 580
DB 353 TGGCAATCCGCACTGACGAAAGATCTGAAATGACCCGCTTTTGTGATGACGCTG 412
QY 581 AAGAAATCCGCTTACCTCATGCTTCAAGAGATCTGAAATGACCCGCTTTTGTGATGACGCTG 640
DB 413 AAGAAATCCGCTTACCTCATGCTTCAAGAGATCTGAAATGACCCGCTTTTGTGATGACGCTG 472
QY 641 TACTGCGAAGCTTCCGCTTACCTCATGCTTCAAGAGATCTGAAATGACCCGCTTTTGTGATGACGCTG 700
DB 473 TACTGCGAAGCTTCCGCTTACCTCATGCTTCAAGAGATCTGAAATGACCCGCTTTTGTGATGACGCTG 532
QY 701 AAGAGCGAAGCGCAAGAGATCTGAAATGACCCGCTTTTGTGATGACGCTG 760
DB 533 GCGCTTACGCTCATGCTTCAAGAGATCTGAAATGACCCGCTTTTGTGATGACGCTG 589
QY 761 AATGATGACGACCTGCTTTTGTGCTGATGCTTAAACCGGCTGCTTAAACGATG 820
DB 590 AATGATGACGACCTGCTTTTGTGCTGATGCTTAAACCGGCTGCTTAAACGATG 649
QY 821 AATGATGACGACCTGCTTTTGTGCTGATGCTTAAACCGGCTGCTTAAACGATG 871
DB 650 AATGATGACGACCTGCTTTTGTGCTGATGCTTAAACCGGCTGCTTAAACGATG 709
QY 872 TTAAGTAAATGCTGAAAGAGTATGATTTGCTCGAAAGCTTGAATTTGAGTAC 931
DB 710 CTGGAATATTTAGTGAAGCAAGGTTATGACCTGTTGAGGCTTGAATTTGAGTAC 769
QY 932 CTTCCTTGGCTATCGAGTTGATCTCAAGAGCTCCGCTTGAATTTGAGTAC 991
DB 770 CTTCCTTGGCTATCGAGTTGATCTCAAGAGCTCCGCTTGAATTTGAGTAC 829
QY 992 CCAAGGTAACCGGCTTGTATCCGAGTATATCCGAACCGGTAATCAACCGGCTAT 1051
DB 830 CCAAGGTAACCGGCTTGTATCCGAGTATATCCGAACCGGTAATCAACCGGCTAT 889
QY 1052 TTGCTGATTTGCTGCAAGTTTGTCTCCCTCCAGTGTGAGTAAATTTAGTAC 1111
DB 890 ACGATGCTGATTTGCTGCAAGTTTGTCTCCCTCCAGTGTGAGTAAATTTAGTAC 949
QY 1112 CCGAGCAATTCGCGCTTCTTTGAGTACCAATTTATTTGATTAATTTCTTAAT 1171
DB 950 TCCGACATGATCGCTGATCTTTG----- 972
QY 1172 TATATTTGTAATAATTTGATTAATTTGATTAATAATGAAATTTGAGTAC 1231
DB 973 ----- 972
QY 1232 GTGACGCTTGTGAAATTAATAAGCAATTAAGGTTTGTGATCAATCAATTAATTA 1291
DB 973 ----- 972

DB 973 ----- 972
QY 1352 AATGAGAGATGATTTTCAAGAGCAACAGACAGTGTGCTTATGAGTACCT 1411
DB 973 -----GAAATGATTTTCAAGAGCAACAGACAGTGTGCTTATGAGTACCT 1026
QY 1412 CGTAGAGATGCTTCAATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1471
DB 1027 CGTAGAGATGCTTCAATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
QY 1472 AGTGGGAAATCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 1531
DB 1087 TGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1146
QY 1532 TGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1591
DB 1147 GGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1206
QY 1592 GGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1651
DB 1207 GGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1266
QY 1652 GGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1711
DB 1267 GGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
QY 1712 GGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1771
DB 1327 GGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1386
QY 1772 GGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1831
DB 1387 GGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
QY 1832 GGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1891
DB 1447 GGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1506
QY 1892 GGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1951
DB 1507 GGTGGAAGAGCAAGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1566
QY 1952 TAAATTTGCGGCTAGGCGCAATTAATAAAG 1983
DB 1567 TAAATTTGCGGCTAGGCGCAATTAATAAAG 1598

RESULT 2
US-08-560-398-3
Sequence 3, Application US/08560398
Patent No. 597082
GENERAL INFORMATION:
APPLICANT: O'Neill, Sharmen
TITLE OF INVENTION: Nadeau, Jeanette
TITLE OF INVENTION: Ovale-Specific Gene Expression
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,398

Tue Jan 20 17:43:21 2004

us-10-022-025a-2.rge

Page 14

IIABMHPHVSQVQEBDVAVGAARVAEDDVAVMTYLPVAVVEVLRHPPELIS
MARLSTINDTIDYHYPAGTATMAYNTALICDHPHYWQDLEPMERPTTGGADEPSI
IGSDPLIAPFGGRBA.CPBGKTGNATVNFVMSLHREBWPSPDEKVDLTVLKLSS
EMANPLTVKVRPRG
BASE COUNT 369 a 457 c 407 g 378 t
ORIGIN

Query Match 61.9% Score 651.6; DB 8; Length 1611;
Best Local Similarity 66.5% Pred. No. 1.1e-186;
Matches 1025; Conservation 0; Mismatches 484; Indels 33; Gaps 5;

61 AGCCAAACCACTTGGCTTCTCCCTGCGCGCTGACATATCTGGCTGCCATATCT 120
74 ACCCAAGAAACCTTGATGGGCTCTTTGATCATGGGCTCACTCTGTTAACATGACT 133
121 CTCTTCTTATGACCTATCCCGGTGACCTGTTGGGGAAATACCTCTTGGCGGTTA 180
134 TTCTATTACTGTCAACCCCGGTGTCTGCTGGGGCAATGACTAACCC-----TA 186
181 ATATCGGCTTATCAAAAACCGAAAGCTTATTCGGGTCCAAAGGCTTCCCTTGGTT 240
187 CTCCTCCCTCTTCA-----TCATTCGGGTCCAAAGGCTTCCCTTATTT 235
241 GGAAGCATGTCATGTCATGACGATCTAGCTCAACGAAATCGCTGATGACCTGAG 300
236 GGAAGCATGAGGCTTATGATTC-----CTGGCCCATCACTGATGACGCGGCGCC 292
301 AATTCGAGCAAGAGGCTCATGCTTCACTTGAAGAGAGATGCTGATCTGACG 360
293 ACATGCAAGAGCAAGGCTCATGCTTCACTTGAAGAGAGATGCTGATCTGACG 352
361 TGCAATCCCGCATGAGGAAAGATCTGATAGCGCGGTTTCTGATGACGCGGT 420
353 TGCAATCCCGCATGAGGAAAGATCTGATAGCGCGGTTTCTGATGACGCGGT 412
421 AAGAGATGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
413 AAGAGATGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 472
481 TACTGGGAGAGCTTCCCGGTATGCTGACCAATCTCTTATGATCAAAACCAATG 540
473 TACTGGGAGAGCTTCCCGGTATGCTGACCAATCTCTTATGATCAAAACCAATG 532
541 AAGAGCAAGAGCAAGAGATGATCTGACCAATCTCTTATGATCAAAACCAATG 600
533 GCTTCTGAGCTCAAGCTCTCAATGCGCGCCCAATGATGATCAATCT--AAATAC 589
601 AGTATGTAAGAACCTCTGTTTGTGCTGATGCTTAAAGCGCGTCTGATCAACATG 660
590 AAGGCGCAACGAGCTTATGCTTGTGCAAGTGTGAAAGGCTTGTGCTGATCAATG 649
661 ATGCTCTGATGCTGCAAGAGATGATGATGATGATGATGATGATGATGATGATG 711
650 ATGCTCTGATGCTGCAAGAGATGATGATGATGATGATGATGATGATGATGATG 709
712 TTAAGTGAATGCTGCAAGAGATGATGATGATGATGATGATGATGATGATGATG 771
710 CTGGAATATGATGCAAGAGATGATGATGATGATGATGATGATGATGATGATG 769
772 CTCTCTGCTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 831
770 CTCTCTGCTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 829
832 CCAAGATGAACCGGTTGTATCCCGATTAATCCGAACCGTATCAACCGGTAT 891
830 CCAAGATGAACCGGTTGTATCCCGATTAATCCGAACCGTATCAACCGGTAT 889
892 TTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
890 ACCAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 949
952 CCGAATATGCTGCTTCTTGGAGATGATGATGATGATGATGATGATGATGATGAT 1011

Db 950 TCCGACATGATGCTGATCTTTGGGAATGATATTCAGAGAAACGACACGATGCGTT 1009
Qy 1012 TTAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1071
Db 1010 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1069
Qy 1072 AACGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1131
Db 1070 GAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1129
Qy 1132 CTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1191
Db 1130 ATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1189
Qy 1192 CTATCATGAGGCGGTTTGGCCATTAACAGACAGATGATGATGATGATGATG 1251
Db 1190 CTATCATGAGGCGGTTTGGCCATTAACAGACAGATGATGATGATGATGATG 1249
Qy 1252 GGAACCAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1311
Db 1250 GGAACCAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 1309
Qy 1312 CCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1371
Db 1310 CAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1369
Qy 1372 CTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1431
Db 1370 CTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1429
Qy 1432 AATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1491
Db 1430 ACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1489
Qy 1492 GGAACCAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 1551
Db 1490 GGAACCAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 1549
Qy 1552 GCTATCT 1593
Db 1550 GCTATCT 1591

RESULT 9
ATH552847/c 740 bp DNA linear PLN 29-MAR-2003
LOCUS Arabidopsis thaliana T-DNA flanking sequence, left border, clone 345B11.
DEFINITION
ACCESSION AJ552847
VERSION AJ552847.1 GI:29368998
KEYWORDS left border; T-DNA flanking sequence.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1
Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Gruard, C., DeRose, R., Pelletier, G., Lepoint, L., Caboche, M., and Lecharny, A. T-DNA integration into the Arabidopsis genome depends on sequences of pre-integration sites
EMBO Rep. 3 (12), 1152-1157 (2002)
TITILE
JOURNAL MEDLINE
PUBMED 22363535
REFERENCE 2 (bases 1 to 740)
AUTHORS Balzerque, S.
TITILE Direct Submision
JOURNAL Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 00:02:34 ; Search time 5880 Seconds
(without alignment)

11083.169 Million cell updates/sec

Title: US-10-022-025a-2

Perfect score: 1593
Sequence: 1 atggctacgaactcgaag.....tgcgcgcgacgacgagttaa 1593

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 segs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Genbank1:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1360	85.4	55870	8 AC005819	AC005819 Arabidops
C 2	1360	85.4	55870	8 AC006418	AC006418 Arabidops
C 3	976	61.3	1605	8 BT008891	BT008891 Arabidops
C 4	976	61.3	1605	8 AY062748	AY062748 Arabidops
C 5	976	61.3	1605	8 AB036059	AB036059 Arabidops
C 6	975.6	61.2	1878	8 AY086928	AY086928 Arabidops
C 7	794	49.8	142001	8 ATE21F14	ATE21F14 Arabidops
C 8	651.6	40.9	1611	8 ATE22463	ATE22463 Glycine m
C 9	575.8	36.1	740	8 ATH552847	ATH552847 Arabidops
C 10	474.4	29.8	86436	8 AC007323	AC007323 Arabidops
C 11	474.4	29.8	104001	8 AC023628	AC023628 Arabidops
C 12	355	22.3	1799	6 ARO70567	ARO70567 Arabidops
C 13	355	22.3	1799	6 PSU34744	PSU34744 Arabidops
C 14	348.2	21.9	1990	8 AF049067	AF049067 Pinus rad
C 15	336.2	21.1	565	11 AL772960	AL772960 Arabidops
C 16	295	18.5	66770	8 AP004704	AP004704 Arabidops
C 17	289	18.1	1712	8 BT000225	BT000225 Arabidops
C 18	289	18.1	2158	8 AY136401	AY136401 Arabidops
C 19	260.4	16.3	504	8 ATH553770	ATH553770 Arabidops
C 20	239	15.0	108387	8 AC016662	AC016662 Arabidops
C 21	237.4	14.9	2087	8 MZCP450A	MZCP450A Arabidops
C 22	236.4	14.8	310	11 AL772952	AL772952 Arabidops
C 23	228.4	14.3	114477	2 AP005682	AP005682 Oryza sat
C 24	218.2	13.7	1668	8 AB096259	AB096259 Oryza sat
C 25	205.6	12.9	122218	8 AC099399	AC099399 Oryza sat
C 26	205.6	12.9	138831	8 AC140005	AC140005 Oryza sat
C 27	198.4	12.5	79054	8 P21F23	P21F23 Arabidops
C 28	194.6	12.2	471	11 AL772558	AL772558 Arabidops
C 29	174.6	11.0	82390	8 AB016893	AB016893 Arabidops
C 30	173.2	10.9	153247	8 AC137992	AC137992 Oryza sat
C 31	155.6	9.8	124573	8 AP005175	AP005175 Oryza sat
C 32	144.2	9.1	1704	8 AF022458	AF022458 Glycine m
C 33	143.6	9.0	169270	8 AC083943	AC083943 Genomic s
C 34	143.6	9.0	324367	8 AE017091	AE017091 Oryza sat
C 35	141.8	8.9	1789	6 AX028817	AX028817 Sequence
C 36	141.8	8.9	1789	6 AR155332	AR155332 Petunia x
C 37	135.6	8.5	1667	6 AX028838	AX028838 Sequence
C 38	135.6	8.5	1764	6 AB057673	AB057673 Torenia h
C 39	134.6	8.4	1545	8 AY082612	AY082612 Ocimum ba
C 40	134.4	8.4	39601	8 AC079678	AC079678 Arabidops
C 41	133	8.3	1554	8 AY082611	AY082611 Ocimum ba
C 42	131.8	8.3	1542	8 AF271649	AF271649 Arabidops
C 43	131.8	8.3	1542	8 AF271650	AF271650 Arabidops
C 44	131.8	8.3	1542	8 AF271651	AF271651 Arabidops
C 45	130.2	8.2	1748	8 AF313491	AF313491 Matchiola

ALIGNMENTS

RESULT 1
AC005819/c
LOCUS
DEFINITION
AC005819
Arabidopsis thaliana chromosome 2 clone T3A4 map C1C0C03, complete
sequence.
AC005819
VERSION
AC005819.3
GI:20197448
KEYWORDS
HTG.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidops.
REFERENCE
1 (bases 1 to 55870)
Rounsley,S.D., Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Mason,T.M.,

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

Shen, M., Romung, C.M., Fraser, C.M., Somerville, C.R. and Venter, J.C.
Unpublished
2 (bases 1 to 55870)
Lin, X.
Direct Submission
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
3 (bases 1 to 55870)
Town, C.D. and Kaul, S.
Direct Submission
Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, cdcrom@igrr.org
On Apr 18, 2002 this sequence version replaced gi:559489.
location/Qualifiers
1..55870

	misc_feature
	1..14747 /note=overlap with BAC clone F13A10 (AC006418:70079..84825)."
gene	92..1276 /gene="ATCg46630" /note="synonym: T3A4.1" <92..1276 /gene="ATc2g46630" 92..1276 /gene="Atcg46630" /codon_start=1 /product="putative extensin" /protein_id="AAC69930.1" /db_xref="GI:3811447" /translation="MSNIRPWRLLSIAAPAAAGSSDPPPOCPPTSSSLVWPAPK PSPPRROPSPROPPSPPOCCOPCTPRCKAPLPSPCSRSPYSPRRHRS BPMPKPAPTPPPPPRSSSTISPSPEKEQELLPKRPNPDPASHSSTSESVKTR SPGSESHNRARAPRVLASFTPLSHSETTKQKLTAETIQTHRHNNQRN HDVNQNHNQNSHTSNQNNQGNPKMRQPSSDSINISTVTITAEKNGAVM EILRSPGMKTGGSGTSHSRYHGEGKRGLQSSSDSGDEKKTKTNVKPNKC SNPMKAFAFNNSWOMINNISIYNSTAISHDPGVHLKISRPGSDMGFFHYSDYGNNGY TN"
repeat_region	723..810 /rpt_family="(CACCAT)n"
repeat_region	1345..1382 /rpt_family="AT_rich" complement(1304..1325) /rpt_family="AT_rich" 3740..5558 /gene="Atc2g46640" /note="synonym: T3A4.2; predicted by genscan and genefinder."
repeat_region	join(<3740..4072,4525..4839,5083..5215,5407..5558) /gene="ATc2g46640" /codon_start=1 /product="hypothetical protein" /protein_id="AAC69921.1" /db_xref="GI:3811438" /translation="WKHEIDVRNKREKKGNEGISEIKRTAILDOVLVDALDNWEK DGVLITGFEDTLNKEDEDMDCECGSLDYVIDGSIIKVNDOSDLIENVEKK FYDHBEHVLDIVDFSGISKLAETPAVLTAAEALETPOKRTTIALFEMDDHKDVG HCCKBNPDNLGEVEFKYONGSKLSKFSPYKKLVMSKSXEENDLRPIKKRKQMI KRMKKIHVEDVATLKADVPKYRNCALEASHHLKIQQBCIGIMWTMYMCFILL TNTKTLKLPVVNVHLYFYVYCFLITYINAS"
CDS	6216..7156 /gene="Atc2g46650" /note="synonym: T3A4.3; supported by full length cDNA: Genes-3743"
RNA	join(6216..6447,6552..6618,6704..7156)
gene	
gene	

CDS

/gene="At2g46650"
join(6370, .6447, 6552, .6618, 6704, .6957)
/gene="At2g46650"
/codon_start=1
/product="putative cytochrome b5"
/translacion="NANISFHDVAKRCKNCDCWIIHGRKYITSTDFEHPGQDNYL
LATGADALIDEVDNSHKDAELMKRYCIGDVDSIVPVTOYIIPWEKHSIAETT
KERSGKKLTIVYLPILTIAGVAFLRPFNKN"
repeat_region
repeat_region
repeat_region
gene
TRNA
CDS

/note="synonym: T3A4_4"
complement(join(-8165, .8782, 9006, .>9980))
/gene="At2g46660"
complement(join(8165, .8782, 9006, .9980))
/gene="At2g46660"
/codon_start=1
/product="putative cytochrome P450"
/protein_id="AAC69922.1"
/db_xref="GI:3831440"
/translation="MATLKSSLYTLFALLSKCVLSGTNIAPSLATVIWLAIPLFM
TYGGAWAKYIFGLRLISGSYKTGVTPGKGFPILVGWSIMLSSTLAHRIRADAAEF
GARLLARFIISTRTIVICNPDAKEHLINSPEAFPKVESAYSLMEFRAIGFAFHV
YATRIIFRINSLTFSTKOIRAEATFORVISQVARELEKOSNECFVRELTKASLN
NMGCSPGOCEYLEKNQHELEEMVEEGYDLGLTWTDHLFWLSFFDFORLSCSTL
VPKNSFVRRIISEHNQGTDPIDPFDVTYLISHGSDKLSDPIIAVLAMEIPEGITLV
PVAILMIILABRVLDHPMDSOTVONEDOVVGSRALDESDIASLPYLAIVAEVLRHL
WATLLSMARLAIIDTVLIDGRLVPAGTTAVNNMAVSHPDHVVDPLEFEKERVAKE
GREYSVGSODLRIAPRGSGRLCPGNLGFTTVMWTAMLHERMGPSDNGNVDS
EKRLSCMANLPKLRRRRS"
10124, .10180
/rpt_family="Ar_rich"
complement(14979, .15002)
/rpt_family="(A)n"
15546, .15568
/rpt_family="Ar_rich"
17214, .17366
/rpt_family="Ar_rich"
17478, .17503
/rpt_family="(A)n"
complement(17523, .17543)
/rpt_family="Ar_rich"
complement(17647, .17688)
/rpt_family="Ar_rich"
complement(17894, .17894)
/rpt_family="Ar_rich"
complement(18001, .18060)
/rpt_family="Ar_rich"
complement(19153, .19796)
/gene="At2g46670"
/note="synonym: T3A4_5; predicted by genescan and
genefinder"
complement(join(<19152, .19310, 19344, .>19796))
/gene="At2g46670"
complement(join(19152, .19310, 19344, .19796))
/gene="At2g46670"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC69924.1"
/db_xref="GI:3831441"
/translation="MYTCRRREESKAERAVNALRESGEPKPTESHKELRYRBD
QGASTTSNQENTIGSSVSFRNKYLOJSTVTYNQGDSPPIVENRKAKSKEVAGSSOS
DSINGAGSSSTRKEBSKAKQSHSQSEALMKFRLLKRCORCFDKHLKIQDT
DSIEWENKLQVRYSRKLEQGFVKGGVFRTVNDSTWS"
19989, .21343
/note="molecular marker Athb7"

Gene	complement (20037 . 21512) /gene="At2g46680" /note="synonym: 13A4.6; identical to SP:P46897; supported by cDNA: gi_15027938_gd_At045826.1" complement (join(20037 . 20039,20111 . 20756,20977 . >21512) /gene="At2g46680" complement (join(20340 . 20756,20977 . 21336)) /gene="At2g46680"
CDS	/codon_start=1 /product="homeodomain transcription factor (ATHB-7)" /protein_id="AAC69925.1" /db_xref="GI:3831442" /translation="MTGGGEGSPAMWSABPFLTMKQKKKSNHNKNNORRFSDEQIKSL EMPFSSERLRPRKKQVLARHILGIPROVALIWNQRRARRKSKOLETEYNILROYND LASQFSLKKRKKQLVSELQRLKATQTKQGEORSCDQAKVNLSSHHSEEN RRRKEEYRPFEMWKDDKGHGVCMDHEDDDGSSNITKATYGGFGPEPDILNM IVRPADSLTRSSMDRGKSDTTLIDQSSNNYPMDWFS"
repeat_region	complement (26336 . 22664) /rpt_family="AT_rich"
repeat_region	23532 . 23571 /rpt_family="AT_rich"
repeat_region	23760 . 23783 /rpt_family="AT_rich"
repeat_region	24595 . 24656 /rpt_family="(CAAA)n"
repeat_region	complement (25135 . 25155) /rpt_family="AT_rich"
repeat_region	25903 . 25923 /rpt_family="AT_rich"
repeat_region	complement (26667 . 26771) /rpt_family="AT_rich"
repeat_region	complement (26870 . 26901) /rpt_family="R:ATREP1 ATREP1 ATREP1 dispersed repeat from GS D86608 - a consensus"
repeat_region	30015 . 30038 /rpt_family="(PAAA)n"
repeat_region	complement (31911 . 31933) /rpt_family="AT_rich"
repeat_region	complement (32447 . 33466) /rpt_family="AT_rich"
Gene	35278 . 35832

Query Match	85.4%	Score 1360;	DB 8;	Length 55870;
Best Local Similarity	87.7%	Pred. No. 0;		
Matches 1593;	Conservative	0;	Mismatches 0;	Indels 223;
				Gaps 1;

QY	1	ATGGCTACGAAACCTCGAAAGCTCCTTATCTTTGACCTTTGTCCTCAATGAGAGGTTCTA	60
Db	9980	ATGGCTACGAAACCTCGAAAGCTCCTTATCTTTGACCTTTGTCCTCAATGAGAGGTTCTA	9921
QY	61	AGCCAAACCAACCTTGGCTTCTCCCTCTCCGCGGTCACAATCATCTGGCTGGCCATATCT	120
Db	9920	AGCCAAACCAACCTTGGCTTCTCCCTCTCCGCGGTCACAATCATCTGGCTGGCCATATCT	9861
QY	121	CTCTTCTTATGACCTATCCGGTGAACCTGTTGGGGAAATACCTTTGGCCCGTTA	180
Db	9860	CTCTTCTTATGACCTATCCGGTGAACCTGTTGGGGAAATACCTTTGGCCCGTTA	9801
QY	181	ATATCCGGTTTATACAAAACCGGAAAGCTTATCCCGGTCCAAAAGGTTCCCTTGGTT	240
Db	9800	ATATCCGGTTTATACAAAACCGGAAAGCTTATCCCGGTCCAAAAGGTTCCCTTGGTT	9741
QY	241	GGAAGCATGTCATCATGTCAAGCACTTAGCTCCACCAATGCTGATGACGTAG	300
Db	9740	GGAAGCATGTCATCATGTCAAGCACTTAGCTCCACCAATGCTGATGACGTAG	9681
QY	301	AAATTCGAGCCAGAGGCTCATGCTTTCAAGCTTAGAGAGACTCGCGTGAATGCTACG	360
Db	9680	AAATTCGAGCCAGAGGCTCATGCTTTCAAGCTTAGAGAGACTCGCGTGAATGCTACG	9621
QY	361	TGCAATCCCGACGTACGGAAGAGATTCTGAATAGCCCGGTTTTGCTGATGACCCGGTT	420
Db	9620	TGCAATCCCGACGTACGGAAGAGATTCTGAATAGCCCGGTTTTGCTGATGACCCGGTT	9561

QY	421	AAAGAAATCGGCTACTGCACTGATGTTTAAACAAGCAATTGGTTTTCGACCAACGGGTGT	480
Db	9560	AAAGAAATCGGCTACTGCACTGATGTTTAAACAAGCAATTGGTTTTCGACCAACGGGTGT	9501
QY	481	TACTGGCGAAGCGTCTCGCCGATCGCTTGGAAACCATCTCTTAACTAACAACAAATCAGA	540
Db	9500	TACTGGCGAAGCGTCTCGCCGATCGCTTGGAAACCATCTCTTAACTAACAACAAATCAGA	9441
QY	541	AGAGCCGAGACGCAACGACGAGTGAATCTCAAGCCAGATGGTTGAAGTTCTTGAAAAACAG	600
Db	9440	AGAGCCGAGACGCAACGAGTGAATCTCAAGCCAGATGGTTGAAGTTCTTGAAAAACAG	9381
QY	601	AGTGTAAACGAACCCGTGTTTGTTCGTAAGTTGCTTAAACCGCGCTCGCTTAAACAATG	660
Db	9380	AGTGTAAACGAACCCGTGTTTGTTCGTAAGTTGCTTAAACCGCGCTCGCTTAAACAATG	9321
QY	661	ATGTGCTCTGTATTCGACACAAGAGTATGAGCTTGAAAAAACAATGTGAGTAGCTGAA	720
Db	9320	ATGTGCTCTGTATTCGACACAAGAGTATGAGCTTGAAAAAACAATGTGAGTAGCTGAA	9261
QY	721	ATGTGCGAAGAGGTATGATTTGCTCGAAGCTGTGAATTGACTGATCACTTCTCTGG	780
Db	9260	ATGTGCGAAGAGGTATGATTTGCTCGAAGCTGTGAATTGACTGATCACTTCTCTGG	9201
QY	781	CTATCGGAGTTTGAATCTCTCAAAAGCTCCGGTCTAGATGTCCACACTCGACCAAAAGTA	840
Db	9200	CTATCGGAGTTTGAATCTCTCAAAAGCTCCGGTCTAGATGTCCACACTCGACCAAAAGTA	9141
QY	841	AACCGGTTGTATCCCGGATTAATCCGAAACCCGTATCAAAACCGGTGATTTGGCTCGT	900
Db	9140	AACCGGTTGTATCCCGGATTAATCCGAAACCCGTATCAAAACCGGTGATTTGGCTCGT	9081
QY	901	GATTTTCGTCGACGTTTTCGCTCTCCCTCCAGTGTCAATGTAATTAATCCGACCCGACATA	960
Db	9080	GATTTTCGTCGACGTTTTCGCTCTCCCTCCAGTGTCAATGTAATTAATCCGACCCGACATA	9021
QY	961	ATCGCCGTTCTTGG-----	974
Db	9020	ATCGCCGTTCTTGGGTAATGACACACATTTATTTGATTAATTAATCTTAAATTAATTTGT	8961
QY	975	-----	974
Db	8960	TGAAAATTCGTTAGATTAATTAATTAACATGAAATTGGAACCTCAATGTGACGTGT	8901
QY	975	-----	974
Db	8900	TGTGGAATTAATTAAGCATTAAGAAATTTTGTGTAATCAATCAATTAAGTAATTTAGAT	8841
QY	975	-----GGA 977	974
Db	8840	TTTATTAACAGTTTCTATATAAAAGTAATAAAGTACATTAATTTTGTAGTATTAATTAACGA	8781
QY	978	GATGATATTCAGAGGAAACGACACAGTTGCGCTCTTAATCGAGTGAATCCTCGCTAGGAT	1037
Db	8780	GATGATATTCAGAGGAAACGACACAGTTGCGCTCTTAATCGAGTGAATCCTCGCTAGGAT	8721
QY	1038	GGTCCCTTATCCAGATATGCAATCAACGCTACAAACGAGCTGATCAAGTAGTCGGGAA	1097
Db	8720	GGTCCCTTATCCAGATATGCAATCAACGCTACAAACGAGCTGATCAAGTAGTCGGGAA	8661
QY	1098	ATCAAGAGCCCTAGATGAATCTGACTTGGCTTCACTTCATATCTAAGCGTTGTGGTGA	1157
Db	8660	ATCAAGAGCCCTAGATGAATCTGACTTGGCTTCACTTCATATCTAAGCGTTGTGGTGA	8601
QY	1158	AGAAAGTAATGAGGCTTCACTCTCAAGGCCCACTTCTATCATGAGGCCGCTTGTGCCATAAC	1217
Db	8600	AGAAAGTAATGAGGCTTCACTCTCAAGGCCCACTTCTATCATGAGGCCGCTTGTGCCATAAC	8541
QY	1218	AGACACGATCGTTGATGTCGTCTTGTTCGGCAGAGGACACACAGAAATGGTGAACATG	1277
Db	8540	AGACACGATCGTTGATGTCGTCTTGTTCGGCAGAGGACACACAGAAATGGTGAACATG	8481

Oy	1278	GCGCGTATCGCATGATCCACACGCTGGAGTTGATCCCTTGAGAGTTAAACTGTAGAGATT	1337
Db	8480	GCGCGTATCGCATGATCCACACGCTGGAGTTGATCCCTTGAGAGTTAAACTGTAGAGATT	8421
Oy	1338	CGTGGCAAAAGAAGAGTGAAGTGAAGTTTCGGTTCTTGGGTGGATTTGAGACTTGCACC	1397
Db	8420	CGTGGCAAAAGAAGAGTGAAGTGAAGTTTCGGTTCTTGGGTGGATTTGAGACTTGCACC	8361
Oy	1398	TTTTCGGGTGGGTGGTGGATTTTGGATCCCCGGGAAAGATCTTGTTTTACTACGTTATGTT	1457
Db	8360	TTTTCGGGTGGGTGGTGGATTTTGGATCCCCGGGAAAGATCTTGTTTTACTACGTTATGTT	8301
Oy	1458	TTCGACGCCGATGATGTTTACATGATGATTGAAATGCGGACCGTCCGATGATAAGCGCGTTGA	1517
Db	8300	TTCGACGCCGATGATGTTTACATGATGATTGAAATGCGGACCGTCCGATGATAAGCGCGTTGA	8241
Oy	1518	CTTATCTCAGAGAACCTGAGAGCTTTCTTGCGAATGGCTATCTCTCTCGCTTAATTCG	1577
Db	8240	CTTATCTCAGAGAACCTGAGAGCTTTCTTGCGAATGGCTATCTCTCTCGCTTAATTCG	8181
Oy	1578	CCGTAGCGCGCAGTTAA	1593
Db	8180	CCGTAGCGCGCAGTTAA	8165
RESULT 2			
AC006418/c			
LOCUS	AC006418	84825 bp	DNA linear PLN 11-MAR-2002
DEFINITION	Arabidopsis thaliana chromosome 2 clone F13A10 map C1C06C03,		
ACCESSION	AC006418		
VERSION	AC006418.4	GI:20197771	
KEYWORDS	HTG.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	SpERMatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.		
	1 (bases 1 to 84825)		
	Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Shen,M., Vanaken,S.E.,		
	Barnstead,M.E., Mason,T.M., Bowman,C.L., Rensing,C.M.,		
	Benito,M.-I., Carrera,A.J., Creasy,T.H., Buehl,C.R., Town,C.D.,		
	Nierman,W.C., Fraser,C.M. and Venter,J.C.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 84825)		
AUTHORS	Lin,X.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-MAR-2002) The Institute for Genomic Research, 9712		
REFERENCE	Medical Center Dr., Rockville, MD 20850, USA		
AUTHORS	3 (bases 1 to 84825)		
TITLES	Town,C.D. and Kaul,S.		
JOURNAL	Direct Submission		
COMMENT	Submitted (27-FEB-2002) The Institute for Genomic Research, 9712		
FEATURES	Medical Center Dr., Rockville, MD 20850, USA, cdowm@igrr.org		
	On Apr 18, 2002 this sequence version replaced gi:5598561.		
source	Location/Qualifiers		
	1..84825		
	/organism="Arabidopsis thaliana"		
	/mol_type="genomic DNA"		
	/cultiVar="Columbia"		
	/CD_xref="taxon:3702"		
	/chromosome="2"		
	/map="C1C06C03"		
	/clone="F13A10"		
	1..36153		
	/note="overlaid with BAC clone F1C10		
	(AC006526:59064..95214)."		
	1..36153		
	/note="overlaid with BAC clone F1C10		
	(AC006526:59064..95214)."		
	complement(1047..3029)		
	/gene="At2g46480"		
	/note="synonym: F13A10.1; predicted by genScan"		
misc_feature			
misc_feature			
misc_feature			
gene			

RNA
 complement(join<1047. .1322,1409. .1518,1595. .1597,
 2074. .2431,2503. .2828,2916. .3029))
 /gene="At2g46480"
 CDS
 complement(join(1047. .1322,1409. .1518,1595. .1597,
 2074. .2431,2503. .2828,2916. .3029))
 /gene="At2g46480"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAD20159.1"
 /db_xref="GI:4415923"
 /translation="MTDACCILKQNEEDKRVRFQHGWTIGAFNDTEMLHRSLOEK
 RLEBANLMDSDLOKLETAAMARSRSVDSAPLGNVTIWNQERYRGSFEDMLQAD
 QILNARYVSGAKTNKLALHOEIEIOTMLAAEESTDIDQCRVLDSDRGVQLA
 RAHQHIECKLITNKLRLAMQIOTVEDELENOTITFLTOLASRALPAHICLITRLNL
 EYHLIDPLNKRFFPRENLENPRLKYLHYVALFEDNLAAVAVVNSIVNANQDSRVYFHLV
 TDKNLFGMSNMFLINPGEATTHVQCPPEPTNINSFVSLDQSSAAKKRYFPTA
 RSEVSGSENLKTRIFRINSMNLNHLFYTFRIFPLEKILFPPDDIVVYQGLTLP
 IDLGVKXSENLDPKFCGMAVGMNIFDLKEMKNNTITFTHQNLENRTILMTGLTP
 PGLTIFVLTQPLQRKMLLGLDGDKIDVKKIERSAVIYHNGHMPWTMGSKTOP
 YTKITNDDHPYIFTCRLFE"
 4060. .4696
 /gene="At2g46490"
 /note="synonym: F13A10.2"
 4060. .4696
 /gene="At2g46490"
 4100. .4504
 /gene="At2g46490"
 /codon_start=1
 /product="unknown protein"
 /protein_id="AAD20160.1"
 /db_xref="GI:4415930"
 /translation="MAADGIFRSIFEGCISGLDSAIRRPYHNKCGCALHDKSGAGAK
 NQNRPRSSGSSSEISFPIKRSSEGNITAMNLFSSSSSSSNQISLSSSSSLSN
 LASDLPVNAAITEPSSRSNKQLRWTIGEDD"
 complement(4483. .4517)
 /rpt_family="CAT'n"
 /rpt_family="Ar_rich"
 complement(6756. .6792)
 /rpt_family="Ar_rich"
 complement(11207. .13602)
 /gene="At2g46500"
 /note="synonym: F13A10.3, similarity to ubiquitin family
 of proteins; supported by cDNA: gi_16930424_gb_Af419566
 1_Af419566"
 complement(join(<11207. .11764,11858. .>13602))
 /gene="At2g46500"
 complement(join(11384. .11764,11858. .13177))
 /gene="At2g46500"
 /codon_start=1
 /product="expressed protein"
 /protein_id="AAD20161.1"
 /db_xref="GI:4415931"
 /translation="MSSAGVALSPVRSBPLIMPLVRANSCIDSYEDDTITLYLTPGS
 VLPMLRVEDSIESVTLRIQSYRGFVVNRQKLVPGRRLASNSNRDYGVSBEGLTL
 LVTKLADVLVLTQTCGKCFHYERGNLTGVKKQISKRGDVPDDEBELLIEG
 KLDSDSLINDICRNDSDVLHVRSAKVRVFPVKNELSLVAEQADKGRKRSK
 VPKKTSIPVVVNSKAKVPLVYKMIQSGDLKSGNSPVRSSEGTGAAFMQGPS
 NKRVGFKTIDEPNANENPQGLPLSPNGEGKKGTXYGEGALREVAAYILDHPSKGS
 KSGEISIEAGVPTPAMTECLHPGNPKIKTKTIGLQWTFTEGDSCEMDGPSFSP
 KVEVHGIIVLDIRLANADHGNILMTKDSKLVAVPLIDREYCIPISEFECTPEWL
 WPAKRPVYAEIOTQYIRLDAEDIDLTFHGMKRPABTAQTLRLSTYLLKGYGRT
 TAFEITICRETLTSKSLVSEEMVEBAQAVLPGTSEAFLEALSDVMDVHLDHSECT
 "

QY	901	GATTTGGTGACGCTTTTGGCTCCGCCCCACANNGTTGAGTAATATTTCCGACCCGGACATA	960
Db	79158	GATTTGGTGACGCTTTTGGCTCTCCCTCCATGGTTTCAGATTAATATTCGACCCGGACATA	79039
QY	961	ATCGCGGTTCTTTG	974
Db	79098	ATCGCGGTTCTTTGGGATGATGACACACATTATTTGATTAATATTTCTTAATATATTTGT	79039
QY	975	-----	974
Db	79038	TGAAAATTCCTTTCGATTTATTTGATTTAAAACATGAAATTTGAGACTCAATGACGTGT	78979
QY	975	-----	974
Db	78978	TGTGGAATATATTAAGCATTAGAAAGTTTTTGTGTCATCAATTAATGTAATTTTAAAT	78919
QY	975	-----GGA	977
Db	78918	TTTATTAACAGTTTCTATATTAATAAGTAATAAAAGTACTATAATTTTGAATTATTAACAGA	78859
QY	978	GATGATATTAACAGAGGAACAGACACAGTTGCGGTCTTAATTCAGTGTGATCCTGCTAGAT	1037
Db	78858	GATGATATTAACAGAGGAACAGACACAGTTGCGGTCTTAATTCAGTGTGATCCTGCTAGAT	78799
QY	1038	GGTCCCTCATCCAGATATGATCAACACAGGTAACAAACAGACTGATCAAGTAGTCGGAA	1097
Db	78798	GGTCCCTCATCCAGATATGATCAACACAGGTAACACAGACTGATCAAGTAGTCGGAA	78739
QY	1098	ATCAAGAGCCTTAGATGAATCTGACTTGCTGCTTCACTTCCATATCTAACCGCTGTGGTGA	1157
Db	78738	ATCAAGAGCCTTAGATGAATCTGACTTGCTGCTTCACTTCCATATCTAACCGCTGTGGTGA	78677
QY	1158	AGAGGTATTTGAGGCTTGATCCTCCAGGCGCACTTCTATCAATGGGCGGTTTGCCATAAC	1217
Db	78678	AGAGGTATTTGAGGCTTGATCCTCCAGGCGCACTTCTATCAATGGGCGGTTTGCCATAAC	78618
QY	1218	AGACACGATCGTTGATGCTGCTCTTTGTTCCGGCACGGGACCAACAGCATGGTGAACATNG	1277
Db	78618	AGACACGATCGTTGATGCTGCTCTTTGTTCCGGCACGGGACCAACAGCATGGTGAACATNG	78555
QY	1278	GGCGGTATTCGATGATCCACACAGTGTGGGTGATCCTTTGGAAGTTTAAACCTGAGAGTT	1337
Db	78558	GGCGGTATTCGATGATCCACACAGTGTGGGTGATCCTTTGGAAGTTTAAACCTGAGAGTT	78499
QY	1338	CGTGGCAAAAGAGGTGAGGTGAGCTTTTCGGTCTTGGGTCGGATTTGAGACTTGAC	1397
Db	78498	CGTGGCAAAAGAGGTGAGGTGAGGTGAGCTTTTCGGTCTTGGGTCGGATTTGAGACTTGAC	78433
QY	1398	TTTGGGGTCGGGTCGTCGATTTTCCCGGGAGAAATCTTGTTTAACTACCGTTATGTT	1457
Db	78438	TTTGGGGTCGGGTCGTCGATTTTCCCGGGAGAAATCTTGTTTAACTACCGTTATGTT	78377
QY	1458	TTGACCGGCGCATGATGATTAACATGATTTGAATGGGACCGTCGATGTAACGGCGTTGA	1517
Db	78378	TTGACCGGCGCATGATGATTAACATGATTTGAATGGGACCGTCGATGTAACGGCGTTGA	78311
QY	1518	CTTATCTAGAAAACGAGGCTTTCTTGCGAGATGGCTAATCCTCTTCTGCTAATTTGCG	1577
Db	78318	CTTATCTAGAAAACGAGGCTTTCTTGCGAGATGGCTAATCCTCTTCTGCTAATTTGCG	7825
QY	1578	CCGTAAGCGCCAGTTAA	1593
Db	78258	CCGTAAGCGCCAGTTAA	78243
<p> RESULT 3 LOCUS BT008891 1605 bp mRNA linear PLN 13-JUN-2003 DEFINITION Arabidopsis thaliana Atcg61880 mRNA, complete cds. ACCESSION BT008891 VERSION BT008891.1 GI:31711947 KEYWORDS Arabidopsis thaliana (thale cress) </p>			

	REFERENCE	TITLE JOURNAL REFERENCE AUTHORS	COMMENT
ORGANISM	Arabidopsis thaliana Arabidops thalliana Eukaryota; Vitellaceae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eucots II; Brassicales; Brassicaceae; Arabidops.		
REFERENCE	1 (bases 1 to 1605)		
AUTHORS	Kim,C.J., Chen,H., Cheuk,R., Shin,P., Bowser,L., Carninci,P., Dale,J.M., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Saitou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.		
TITLE	Arabidopsis ORF clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1605)		
AUTHORS	Kim,C.J., Chen,H., Cheuk,R., Shin,P., Bowser,L., Carninci,P., Dale,J.M., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Saitou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-JUN-2003) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA		
COMMENT	Riken Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Saitou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.		
	The Salk, Stanford, PEBC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Kim,C.J., Chen,H., Cheuk,R., Shin,P., Bowser,L., Chan,M., Chang,C.M., Dale,J.M., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Davis,R.W., Theologis,A., and Ecker,J.R.		
	Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.		
FEATURES	Location/Qualifiers		
SOURCE	1..1605 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="3" /clone="U16612" /note="This clone is in pUNI 51 ecotype: Columbia"		
CDS	1..1605 /note="cytochrome p450 (CYP78A9)" /codon_start=1 /product="At3g61880" /protein_id="AAPE8330.1" /db_xref="GI:31711948" /translat="MAYKDDTSSLALLALKSKGLITOTNLASLLVASLASLSTLSFPWSHSGPWANGYPLARRRQTTVPFGPGLEPFVSMISMTVLNHRCLATAEKRPRERLMASLGRTIVTCNDVAKEILNSVPEADRPVKSAVSLAMNRALIGAPAVCYWMTLRKIASLSPSPKOIKRSQTSRYANAIQVCITKQSNTKGCLFPARDLITVALNNMMCSVGKVEILEEHEEYSSELREIYBEGDLDGLTNIMTDHLPMISEPPQIRSRCSLLVPKRNFTVNIIISHREBOTSDPSDFLVLLDGPKSLDPDI IAVIMIEIFRGDT VAVIIEKLIAWTIIPDICSTVHNELIDQIAGRBAVESVSIVLYTAVAKELRAH PGSEGLSARLATIDTTLIDERRVDCAGCTAMWNMAIALDHWNENPLEKPERFAKE GEFFSVSGDRLRAPFGSGRRVCGRCCNGTLGTYTFMTAILHFEMLTSPSDEKTDL SEKTRISCENANPLAALRLPRSPSV"		
BASE COUNT	378 A 400 C 395 G 432 T		
Query Match	61.3%; Score 96; DB 8; Length 1605;		

Best Local Similarity 78.0%; Pred. No. 1.5e-285;
Matches 1238; Conservative 0; Mismatches 320; Indels 30; Gaps 4;

QY 19 AGCTCTTAATCTTTGGCCCTTTTGTCCAAATGCAAGCGTTTGAAGCCAAACCAACTTGGC 78
Db 22 AGTAGCTTACTTTGGCCCTCTGTCCAAATGAGCCCTCTTACTCAAAACCAATCTTGTCT 81
QY 73 TTCTCCCTCCCGCCGTCACAAATCATCTGCTGCCATCTCTCTTCTTAAGACCTAT 138
Db 82 CTCTCTCTCCGAGCTCCCTAGCTTCTCTGCTCTTCTCTCTCTCTCTCTCTCTCTCT 141
QY 139 CCGGTGAGCTGCTGGGGGAAATACCTCTTGGCCGGTTAAATCCGGTTCAACAA 198
Db 142 CCGGAGAGACCGGATGGGGAAAGTACTTCTCCACCGC-----CGCCGT 186
QY 199 ACCGAAACGTTATTTCCCGTCCAAAGGCTTCCCTTGGTTGAAGCACTCACTCATG 258
Db 187 CAACCAACCGTAGACCCGGCCCAAGAGCTTACTTTTGTGGAAGCATGTCTCATG 246
QY 259 TCAGCACTAGCTCAACGCAAGATGCTGATGCACTGAGAAATTCGAGCCAAAGAG 318
Db 247 TCAAACTCTGCTCAACGCTGCAAGCCGCAACCGCAAGAAATTTAGAGCGAAAG 306
QY 319 CTATGCTTTACGCTTGAAGAGACTGCGGTGATGCTCACTGCAATCCGACGTAGCG 378
Db 307 TTAATGGGTTAGTTGGGAGAACTCGGTGATGCTCACTGCAATCTGATGTAAGT 366
QY 379 AAAGAGATTCGAATAGCCGGTTTGTGCTGATGACCGGTTAAAGATCGGCTTACTCA 438
Db 367 AAAGAGATTCGAATAGCCGGTTTGTGCTGATGACCGGTTAAAGATCGGCTTACTCA 426
QY 439 CTATGTTTAAACAGACCAATGTTTTCACACAGGTTTAACTGAGCGATTCGC 498
Db 427 CTATGTTTAAACAGACCAATGTTTTCACACAGGTTTAACTGAGCGATTCGC 486
QY 499 CGTATCGTTGAACCAATCTTGTAGTAAACAAATCAGAAAGCGAGACGCAAGA 558
Db 487 AAATTCGGCTTAATCATCTTTTCAAGCCGAAACGATTAACGTTCCGAAACGCAAGA 546
QY 559 CGAGTATCTCAAGCCAGATGTTGATTTCTTGAATAACAGATAGTAAAGAA---CCC 615
Db 547 AGCGTATCGGATCAATCATGATGATGATGATGATGATGATGATGATGATGATGAT 606
QY 616 TGTGTTTGTGAGTGTCTTAAACAGCGGCTGCTTAAACAGATGATGATGATGATGAT 675
Db 607 TGTGTTGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666
QY 676 GAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
Db 667 GGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
QY 727 GAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
Db 727 GAGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
QY 787 GAGTGTATCTCAAGACTCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 846
Db 787 GAGTGTATCTCAAGACTCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 846
QY 847 TTTGATCCCGATTAATCCGACACCGGATTAACCGGATTTGCTGCTGATTTCC 906
Db 847 TTTGATCCCGATTAATCCGACACCGGATTAACCGGATTTGCTGCTGATTTCC 906
QY 907 GTGCAAGTTTGTCTCTCCCTCCATGATGATGATGATGATGATGATGATGATGATGATGAT 966
Db 907 GTGCAAGTTTGTCTCTCCCTCCATGATGATGATGATGATGATGATGATGATGATGATGAT 966
QY 967 GTTCTTTGGAGATGATTAATCAAGAGACAGACAGATGCGGTCTTAATGAGTGAATC 1026
Db 967 GTTCTTTGGAGATGATTAATCAAGAGACAGACAGATGCGGTCTTAATGAGTGAATC 1026
QY 1027 CTCCTAGAGATGCTCTTCAATCAGATGATGATGATGATGATGATGATGATGATGATGAT 1086

Db 1027 CTGCTAGAGATGCTCTTCAATCAGATGATGATGATGATGATGATGATGATGATGATGAT 1086
QY 1087 GTAGCCGGGAATCAAGAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1146
Db 1087 ATCGGGAGATCAAGAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1146
QY 1147 GCTGTGTGAAGAAGTATGAGCTTCACTTCAAGCCCACTTCAATGAGGCGCCGT 1206
Db 1147 GCTGTGTGAAGAAGTATGAGCTTCACTTCAAGCCCACTTCAATGAGGCGCCGT 1206
QY 1207 TTGGCAATACAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1266
Db 1207 TTGGCAATACAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1266
QY 1267 GTGAACATGTGGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
Db 1267 GTGAACATGTGGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
QY 1327 CCGTGAAGTGTGCGCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1386
Db 1327 CCGTGAAGTGTGCGCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1386
QY 1387 AGACTTGACCTTTGCGGTCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
Db 1387 AGACTTGACCTTTGCGGTCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
QY 1447 ACCGTGATGTTTGAACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1503
Db 1447 ACCGTGATGTTTGAACGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1503
QY 1504 GGTAAACGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1563
Db 1504 GGTAAACGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1563
QY 1564 CCTGCTAAATTCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1591
Db 1564 CCTGCTAAATTCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1591

RESULT 4
AY062748 1824 bp mRNA linear PIN 25-NOV-2001
LOCUS Arabidopsis thaliana cytochrome p450 (CYP78A9) (F21F14.5) mRNA,
DEFINITION complete cds.
ACCESSION AY062748
VERSION AY062748.1 GI:117065143
KEYWORDS F1, CDNA,
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 1824)
Southwick, A., Karlin-Neumann, G., Nguyen, M., Lam, B., Miranda, M.,
Palm, C.O., Bower, L., Jones, T., Banu, J., Carninci, P., Chen, H.,
Chen, R., Chung, M.K., Hayashizaki, Y., Ishida, J., Kamiya, A.,
Kawai, J., Kim, C., Lin, J., Liu, S.X., Narusaka, M., Pham, P.K.,
Sakano, H., Sakurai, T., Saito, M., Seki, M., Shim, P., Yamada, K.,
Shinozaki, K., Ecker, J., Theologis, A. and Davis, R.W.
Direct Submission
Submitted (14-NOV-2001) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arab@sequence.stanford.edu
Riken Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN
Arabidopsis Full-length cDNA"): Seki, M., Narusaka, M., Ishida, J.,
Saito, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.
The Salk, Stanford, PGSC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Nguyen, M.,

QY 1564 CCTGCTAATTTGGCCGCTAGCGCAGTT 1591
 Db 1620 GCTGCTAATTTAGCCGCCAGCGCAGTT 1647

RESULT 5
 AB036059
 LOCUS
 DEFINITION
 AB036059 1902 bp mRNA linear PLN 23-DEC-1999
 Arabidopsis thaliana CYP78A9 mRNA for cytochrome P450, complete cds.

ACCESSION
 AB036059
 VERSION
 Arabidopsis thaliana (thale cress)
 KEYWORDS
 Arabidopsis thaliana
 SOURCE
 Arabidopsis thaliana
 ORGANISM
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
 1 (bases 1 to 1902)
 Ito, T.
 cytochrome P450 (CYP78A9)
 Published Only in Database (1999)
 2 (bases 1 to 1902)
 Ito, T.

AUTHORS
 Submitted Submission
 Technology, Division of Biology, 1200 E. California Blvd, Pasadena, CA 91106, USA (E-mail: itot@cco.caltech.edu, Tel: 626-395-4936)

FEATURES
 source
 1. 1902
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /cultivar="Landsberg erecta"
 /db_xref="taxon:3702"
 /chromosome="3"
 /map="83 cM"
 /clone="28-5"
 1. 1902
 /gene="CYP78A9"
 69. 1673
 /gene="CYP78A9"
 /product="cytochrome P450"
 /protein_id="BA88569.1"
 /db_xref="GI:6633989"
 /translation="MATKLDTSLLALSKSILITQTNLALSLVSLASLISLFP
 WSHGGRAMKRYFLARRQTTVTPGPRGLFVPSGMSLMSNTLAHRCIAAAREPABR
 LMAPSLDETRIVTCNPDAKEILNSVFPADRPKESAYSLMNRRAIGFAPVMT
 LRKIASNHLSPKQIKRSEIORSVIANQIVKLCIKOSNTGICFARDLKTASLNNM
 CSVPGEYLEEEHEEVESELEIYERGVILGLTMDHLPMISEDPORIKRSCNL
 VPKNRVFNRIISDRBQTRDSPDFVLLSLDGPDKSLDPDIILAVLMIRGSDT
 VAVLIEWILAMVLPDIQSTVHNELDQIVSRSAVESDVSLVYLAVKVKRLH
 PPGILSMARLITDITIDGRVPAAGTAVNMWALAHDHWMENLEPEEFVAKB
 GEVFSYLGSLRIAPRGSGRRVPGNLTGTYTTFATLILHEFMTLSDEKTYDL
 SEKRLSCENWPLAKLRPRRSV"

BASE COUNT 507 a 448 c 439 g 508 t

ORIGIN

Query Match 61.3% Score 976; DB 8; Length 1902;
 Best Local Similarity 76.0%; Pred. No. 1.5e-265;
 Matches 1238; Conservative 0; Mismatches 320; Indels 30; Gaps 4;

QY 199 ACCGGAAACGTTATCCCGGTCGAAAGGCTTCCTTGGTTGGAAGCATCTCACTATG 258
 Db 255 CAACACACCGTGAATACCGGCGCAAGAGGCTTACCTTTTGGAGAGCATCTCTCATG 314
 QY 259 TCAAGCACTAGCTACCGACGAAATGCGTATGAGTGAATTCGAGCAAGG 318
 Db 315 TCAACACTCTGGCTACCGCTTGCATACCGCAACCGAGAAATTTAGCCGAAG 374
 QY 319 CTGATGCTTCAAGCTTAGAGAGCTCGTATGCTCACTGCAATCCGAGTACG 378
 Db 375 TTAATGGCGTTAGTTAGTGGAGAACTGCGTATGCTCACTGCAATCCGAGTACG 434
 QY 379 AAAGGATTCGATAGCGCGGTTTTCGATGACCGGTTAAGAAATGGCTTACTCA 438
 Db 435 AAAGGATTCGATAGCGCGGTTTTCGATGACCGGTTAAGAAATGGCTTACTCA 494
 QY 439 CTGATGCTTCAAGCAATGCTTTTGCACACGCGTGTTCCTGCGAACCGTTGCG 498
 Db 455 CTGATGCTTCAAGCAATGCTTTTGCACACGCGTGTTCCTGCGAACCGTTGCG 554
 QY 459 CGTATGCTTGAACCATCTCTTTAGTCAAAACAAATCAGAAAGCCGAGACGACGA 558
 Db 555 AAATGCGGTCATATCATCTTTTCAGCCCGAAACGATTAACGTTCCGAAACGAGAGA 614
 QY 559 CGAGTATCTCAAGCCAGATGAGTTGATTTCTGAAAACAGAGTATGACAA---CCC 615
 Db 615 ACCTGATCGCAATCAATGCTGATGCTCTCAAAACAGATGACAAACGAAAGGCTC 674
 QY 616 TGTGTTGCTGAGTCTTAAACGCGCTGCTTAAACATGATGCTCTGATTC 675
 Db 675 TGTGTTGCTGAGTCTTAAACGCGCTGCTTAAACATGATGCTCTGATTC 734
 QY 676 GGACAAAGTATGAGCTTGAAGAAACATG-----TTGATTCGTAAGTATGCT 726
 Db 735 GGAAAGATTCAGACTTGAAGAAACATGAAAGTATGAGTATGCTGATTCG 794
 QY 727 GAAGAGGTTATGATTTGCTGGAAGTATGAGTATGAGTATGAGTATGAGTATG 786
 Db 795 GAAGAGGTTATGATTTGCTGGAAGTATGAGTATGAGTATGAGTATGAGTATG 854
 QY 787 GAGTTGATCTCAAGACTCGCGCTGATGTTCCACATCGTACCAAGATTAACCG 846
 Db 855 GAGTTGATCTCAAGACTCGCGCTGATGTTCCACATCGTACCAAGATTAACCG 914
 QY 847 TTTGATCCCGGATATATCCGACACCGTATCAACCGGATGATTTGCTGATTC 906
 Db 915 TTTGATCCCGGATATATCCGACACCGTATCAACCGGATGATTTGCTGATTC 974
 QY 907 GTGACGTTTCTCTCCCTCATGCTTGAATTAATATTCGACCCGGAATATCGCC 966
 Db 975 GTTACGTTTCTCTCCCTCATGCTTGAATTAATATTCGACCCGGAATATCGCC 1034
 QY 967 GTTCTTTGGAGATGATTTGAGAGAAACAGACAGTTCGGCTTATCGAGTATC 1026
 Db 1035 GTTCTTTGGAGATGATTTGAGAGAAACAGACAGTTCGGCTTATCGAGTATC 1094
 QY 1027 CTGCTAGGATGATCTTCATCCAGATATGCAATCAAGGTAACAAACGAGTATCA 1086
 Db 1095 CTGCTAGGATGATCTTCATCCAGATATGCAATCAAGGTAACAAACGAGTATCA 1154
 QY 1087 GTAGTGGGAATCAAGGCGCTTGAATGATGATGATGATGATGATGATGATGATG 1146
 Db 1155 ATGTGGGAATCAAGGCGCTTGAATGATGATGATGATGATGATGATGATGATG 1214
 QY 1147 GCTGTGTAAGAAAGATATGAGCTTCACTTCAAGGCGCACTTATCATGAGCCGT 1206
 Db 1215 GCTGTGTAAGAAAGATATGAGCTTCACTTCAAGGCGCACTTATCATGAGCCGT 1274
 QY 1207 TTGGCATACAGACAGATGCTTATGATGATGATGATGATGATGATGATGATGATG 1266
 Db 1275 TTGGCATACAGACAGATGCTTATGATGATGATGATGATGATGATGATGATGATG 1334

QY 1267 GTGAACATGTGGGCGGATTCGATCATCAGCGTGTGGTTGATCTTTGAGTTTAA 1326
 Db 1335 GTGAACATGTGGGCGGATTCGATCATCAGCGTGTGGTTGATCTTTGAGTTTAA 1394
 QY 1327 CCGTGAAGTTCGTGGCAAAAGAGTGAAGGAGTTTGGTTCTTGGTGGATTTG 1386
 Db 1395 CCGGACGTTTGTGACCAAGAGTGAAGTGAAGTTCGTGGTTCTTGGTGGATTTG 1454
 QY 1387 AGACTTGACCTTTGGGTTGGGTTGGTGGATTTGGCCCGGAGAAATCTTGGTTTAA 1446
 Db 1455 AGGCTTGACCGTTGGGTTGGGTTGGTGGATTTGGCCCGGAGAAATCTTGGTTTAA 1514
 QY 1447 ACCGTTAGCTTTGACCGGATTCGATCATCAGCGTGTGGTTGATCTTTGAGTTTAA 1503
 Db 1515 ACCGTTAGCTTTGACCGGATTCGATCATCAGCGTGTGGTTGATCTTTGAGTTTAA 1574
 QY 1504 GGTAAACGCGTTGACTTATCTGAAGAACTGAAGGCTTTCTTGGAGATGCTTAATCTTCT 1563
 Db 1575 GAAAGACCGTTGACTTATCTGAAGAACTGAAGGCTTTCTTGGAGATGCTTAATCTTCT 1634
 QY 1564 CTTGCTTAATTCGCGCGTTGGGCGGAGTT 1591
 Db 1635 GCTGCTTAATTCGCGCGGAGGCGGAGTT 1662

RESULT 6
 AY086928 1878 bp mRNA linear PLN 14-APR-2003
 LOCUS AY086928 Arabidopsis thaliana clone 29661 mRNA, complete sequence.
 DEFINITION AY086928.1 GI:21405652
 ACCESSION FLI CDNA.
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1878)
 Haas,B.J., Volfovsky,N., Town,C.D., Troupkhan,M., Alexandrov,N.,
 Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
 Full-length messenger RNA sequences greatly improve genome
 annotation
 Genome Biol. 3 (6), RESEARCH0029 (2002)

TITLE JOURNAL MEDLINE PubMed
 22088475
 12093376
 2 (bases 1 to 1878)
 Brover,V., Troupkhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
 Feldmann,K.
 Full-length cDNA from Arabidopsis thaliana
 unpublished
 3 (bases 1 to 1878)
 Brover,V., Troupkhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
 Feldmann,K.
 Direct Submission
 Submitted (11-MAR-2002) Ceres, Inc., 3007 Malibu Canyon Road,
 Malibu, CA 90265, USA
 COMMENT
 This clone sequence is one of 5,000 Ceres full-length cDNAs made
 available to TIGR and Genbank. The following quality assessment of
 this set was done by comparison with known proteins: two percent of
 the clones are estimated to be 5'-truncated; less than one percent
 are 3'-truncated; approximately two percent represent alternative
 splice variants, including unspliced introns and spliced exons; one
 percent may contain premature stop codons; five percent may have
 frame shifts in a coding region. A sequence is considered to be
 5'-truncated if it lacks the translation initiation start (ATG). A
 sequence is considered to be 3'-truncated if it lacks the
 C-terminal end of the encoded protein. Please note that these cDNA
 sequences are derived from the WS or LAR ecotypes and therefore
 may contain polymorphisms when compared to sequences from Col-0.
 Genest carried out the library production and sequencing of the
 full-length clones. Ceres, Inc. carried out the clustering of the
 5' sequences, selection of clones, and sequence assembly.
 Location/Qualifiers

source 1.1878
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="29661"
 51..1655
 /codon_start=1
 /product="cytochrome P450 (CYP78A9)"
 /protein_id="AA064492.1"
 /db_xref="GI:21592543"
 /translation="MATKIDTSLLIALSKSLTQTNLALSLIVASLANSLEF
 WSPGGPAMGKTFELRRQTVIPGPRGLPEVGSMSLMSNTLAHRCIALATKFAER
 LMAFSLGETRVIATCPDYAKEILNSPVFADPVEKESYSLMFNPAIGAPGVWRT
 LRIASNLHFSRQIKRSETRSVIANQICLTQSNYKGLCPARDILKIASLNMM
 CSVQKREYELFEHEHEVSEIRLVEBGDLGLTAMTDLHPVLSFEPDRIKSCNL
 VPKVRFPVRIISDRRECRDSRDPVYVLSIDGPKLSDPDIIVATVAVYEVVRLH
 VAVLLEWILARVILHEDIDISTYHNELDQVGSRAVESDVYSIVYLAIVYEVVRLH
 PGPPLSMARLAIITDTLLIDGRVVPGLTAMVMMALADPHYEMPLKPKRPRVAKS
 GEVSEVLSGDLRIAPFGGRVCPGKMLGLTIVFWATLILHEEXLTPSDKTVDL
 SEKRLRSCMANPLAKLRPRRSFV"
 BASE COUNT 501 a 441 c 436 g 499 t 1 others
 ORIGIN

Query Match 61.2%; Score 975.6; DB 8; Length 1878;
 Best Local Similarity 77.9%; Pred. No. 2e-285;
 Matches 1237; Conservative 1; Mismatches 320; Indels 30; Gaps 4;

QY 19 AGCTCCATATCTTGGCCCTTTGTGCAATGAGCGGTTCTAAGCCAAACCACTTGC 78
 Db 72 AGTAGCTACTTTTGGCCCTTGTCCAAATGAGCTCTTAATCAACCACTTGTCT 131
 QY 79 TTCTCCCTCTGCGCCGTCAACATCATCTGAGTCCATATCTTCTTTATGACCTAT 138
 Db 132 CTCTCTCTCTCTTACCTCCCTAGCTTCTCTGCTCTTCTCTTCTTCTGCTCAT 191
 QY 139 CCGGTTGAGACCTGTTGGGGGAAATACCTTTGCGCGGTTATATCCGTTCAACAA 198
 Db 192 CCGGAGAGCCCGCAATGGGAAAGTACTTCTCCACCGC-----CGCGGT 236
 QY 199 ACCGAAACGTTATCCCGGTCCAAAGGCTTCCCTTGGTTGAGACATGTCATG 258
 Db 237 CAACACCGGTATACCCGGCCCAAGAGGCTTACTTTGTGGAAGCATGTCATG 296
 QY 259 TCAAGCATTTACTACCGACGAATCGGTGATGAGCACTGAGAAATTCGAGCCAGAG 318
 Db 297 TCAACACCTGTGCTCACCGTTCAATGACCGCAACCGAGAAATTTAGGCCAGAG 356
 QY 319 CTCATGCGTTTACGTTTGAAGAGACTCGGTATCGTCACTGCAATCCGACGTAGCG 378
 Db 357 TTATAGCGTTTATGTTTGGGAAACCTCGGTATCGTCACTGCAATCCGATGTAGCT 416
 QY 379 AAGAGATTTGTAATAGCCCGGTTTGTGTCATGACACCGGTTAAAGATGGCTACTA 438
 Db 417 AAGAGATTTGTAACAGTCCGGTTTGTGTCATGACACCGGTTAAAGATGATTATTC 476
 QY 439 CTGATGTTTAAACAGAGCAATGGTTTGAACACACGCTGTTTACTGCGACACGCTTGC 498
 Db 477 CTGATGTTTAAACGCTGCTATCGGTTTGTGCTTACCGGCTTTACTGCGACACGCTTGA 536
 QY 499 CGTATCGTTGAAACATCTCTTTAGTCAAAACCAATCGAAGCCGAGAGCAACGA 558
 Db 537 AATATCGGCTTATCATCTTTTACGCGGAAACAGTTTAAAGTTCCGAAAGCAGAGA 596
 QY 559 CGAGTATCTCAACGAGATGTTGATTTCTTGAACAAACAGATGTAACGA---CCC 615
 Db 597 AGCGTATCGGATCAAAATGTAAGTGTCTACAAAACAGATGTAACCAAAAGGCTC 656
 QY 616 TGTGTTGTTGTTGATGTTCTTAAACGGCGTCTTAAACAAATGATGTTCTGTTTC 675
 Db 657 TGTGTTGATGTTGATGTTCTTAAACGGCGTCTTAAACAAATGATGTTCTGTTTC 716
 QY 676 GGAACAAGATGATGTTGAAAAAACATG-----TTGAGTTACGTAATGCTC 726

Db 717 GGAAGAAATACAGAGCTTGAGGAGACATGAAAGAGTAGAGCTGACGTGAATTGGTC 776
 Qy 727 GAAGAAAGTATGATTTGCTCGAAGCTGGAATTGAGTATGATCACTTCTTGGCTATCG 786
 Db 777 GAAGAGGTATGATTTACTCGGTACCTGAATTTGAGCCGATCATCTCCATGGCTCTCT 836
 Qy 787 GAGTTGATCTCTCAAGACTCCGGTCTAGATGTTTCCAGACTCGTACCAAGGTAACCCG 846
 Db 837 GAATTGATCTCTCAAGAAATCCGGTCTAGATGTTTCCAGACTCGTACCAAGGTAACCCG 896
 Qy 847 TTTGATATCCCGAATTATATCCGAACACCGTATCAACCCGGTATTTGGCTCGTATTC 906
 Db 897 TTTTGACCCGGATATCTCTGACACCGGTAACCAAACTGTAACCTGACCTGACCTG 956
 Qy 907 GTGACGTTTGGCTCTCTCCCTCCATGTTTCAAGTAAATATCCGACCCGACATATACCC 966
 Db 957 GTTGACGTAATGCTCTCTCCATGTTTCAAGTAAATATCCGACCCGATATCAATCCG 1016
 Qy 967 GTTCTTGGAGATGATATCAGAGAAACAGACAGTTCGGTCTTAATCGAGTGATC 1026
 Db 1017 GTTCTATGGAGATGATATCAGAGAACTGACAGCTGCTGTTTGAATCGAGTGATC 1076
 Qy 1027 CTCCTAGAGATGATCTCTTCAATCCAGATATGCAATCAACGCTACAAAAGAGCTGATCA 1086
 Db 1077 CTGCTAGAGATGATCTCTTCAATCCAGATATCAATCGAGTTCACATGAGCTGATCA 1136
 Qy 1087 GTATCCGGAAATCAAGACCCCTAGATATCTGAGCTTCACTGCTCATATCTAAAG 1146
 Db 1137 ATCTGGAGCAGATCAAGGCTGTGAAAGCTGACGCTGTTCTAGATATCTAAAG 1196
 Qy 1147 GCTGTGGTGAAGAGATGAGGCTTCACTCTCCAGCCCACTTATCATGAGCCCGT 1206
 Db 1197 GCTGTGGTGAAGAGATGAGGCTTCACTCTCCAGCCCACTTATCATGAGCCCGT 1256
 Qy 1207 TTGGCCATACAGACAGATGTTGATGTTGTTTCTTCCGAGAGGACCAAGCATG 1266
 Db 1257 TTACCATACAGACAGATGTTGATGTTGTTTCTTCCGAGAGGACCAAGCATG 1316
 Qy 1267 GTGAACATGAGGCGGTATCGATGATCACAAGCTGTGAGTGTATCTTGGAGTTTAA 1326
 Db 1317 GTGAACATGAGGCGGTATCGATGATCACAAGCTGTGAGGAGATCTGTTGAGTTTAA 1376
 Qy 1327 CCTGAGAGTTCGTGGCAAAAGAGTGAAGTTCGTTTGGTCTTGGTGGATTTG 1386
 Db 1377 CCCGAACGTTTGTAGCCAAAGAGTGAAGTTCGTTTGGTCTTGGTGGATTTG 1436
 Qy 1387 AGACTGACCTTCCGGTGGGTGCTGAGATTTGCCCGGGAAGATCTTGGTTTACT 1446
 Db 1437 AGGCTTGAACCGTTCGGGTGCTGAGATTTGCCCGGGAAGATCTTGGTTTACT 1496
 Qy 1447 ACCGTTATGTTTGAACCGGATGATGTTTATCATGATTTGAATG--GGACCGTCCGAT 1503
 Db 1497 ACCGTTATGTTTGAACCGGATGATGTTTATCATGATTTGAATGATCCGCTCCGAT 1556
 Qy 1504 GGTACCGGCTGATGATTTGAGAACTGAGAGCTTCTTCCGAGATGAGTATCTCTT 1563
 Db 1557 GAGAAACCGTGTGATTTGAGAACTGAGAGCTTCTTCCGAGATGAGTATCTCTT 1616
 Qy 1564 CCGTCTAAATGCGCGGTAGCGGCGAT 1591
 Db 1617 GCTGCTAAATGCGCGCGGTAGCGGCGAT 1644

RESULT 7
 APT21F14/C
 LOCUS APT21F14 142001 bp DNA linear PLN 02-FEB-2000
 DEFINITION Arabidopsis thaliana DNA chromosome 3, BAC clone F21F14.
 ACCESSION AL138642
 VERSION AL138642.1 GI:6899881
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatozoa; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroside II; Brassicales; Brassicaceae; Arabidopsis.
 1
 Choisme, N., Robert, C., Broetier, P., Wincker, P., Catolico, L.,
 Artiguenave, F., Saurin, W., Weissenbach, J., Mewes, H.W., Lemcke, K.,
 Mayer, K.F.X., Queller, F. and Salanoubat, M.
 Unpublished
 2 (bases 1 to 142001)
 EU Arabidopsis sequencing project.
 Direct Submission
 Submitted (02-FEB-2000) MIPS, at the Max-Planck-Institut fuer
 Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG. E-mail:
 lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project
 Coordinator: Marcel Salanoubat and Francis Queller, Groupement
 d'Interet Public, Centre National de Sequencage - GENOSCOPE, 2 rue
 Gaston Cremieux, BP191, 91006 Evry Cedex, France;
 http://www.genoscope.cns.fr
 Information on performance of analysis and a more detailed
 annotation of this entry and other sequences of chromosomes 3, 4
 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.
 FEATURES
 source
 1..142001
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /varietal="Columbia"
 /db_xref="taxon:3702"
 /chromosome="3"
 1..56180
 /note="overlap to BAC F15G16, please refer also to Acc.
 no. BMBJ:AL132959 for analysis and annotation"
 5113..55887
 /gene="F21F14.10"
 join(5113..55695,55835..55887)
 /gene="F21F14.10"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CB71891.1"
 /db_xref="gi:6899882"
 /translation="WVIVKRCIFRNDYIYLHCLLOTNNHKKOEDWVTKOKNK
 KDTSPDADNETMILITSLQHVLAQRSRLQYSSPVTTQVSLSLMSDDSLRA
 SLPFSEBEPGKPHVHVPSSKCLDLPMLPDEFTKPRTHRGVYLAGWGRWF
 RKVKRQDQAVYVAGSGSYTPESEDDTKTRTKVPMKXCLRKASVPA"
 5113..55695
 /gene="F21F14.10"
 /number=1
 55696..55834
 /gene="F21F14.10"
 /number=1
 55835..55887
 /gene="F21F14.10"
 /number=2
 57967..57993
 /note="27 bp TCT tandem repeat"
 58167..59822
 /gene="F21F14.20"
 join(58167..58190,58872..59376,59461..59822)
 /gene="F21F14.20"
 /codon_start=1
 /product="transcription factor BBFa"
 /protein_id="CB71892.1"
 /db_xref="gi:6899883"
 /translation="MDATRWQGFQEMINVKEMQNISSTNNTPQOFTPTATNTP
 NATASNGSGGANTNNVATWETRKARPOEKNPCRNSTTRCYNNYSLTQPRFCK
 GGRVYTBGSSLRNVGVGSSRKXSSSTPLASPSNKLPLDANPLLSQIPNKNK
 DNTLSFPMQDHHALBELLSNGVSRGKMTPLFGQWMDNSVLSYSGPTWMDY
 KQNNNTLSSTIHQIGHTINSNRQADNDNDMGARVLPFEDMKELSTTQEK
 SHGNTYKNGMFSNTGSSW"
 58167..58190
 /gene="F21F14.20"
 /number=1
 58191..58871
 /gene="F21F14.20"
 /number=1
 exon
 intron

```

repeat_unit      58253..58276
                  /gene="F21F14.20"
                  /note="24 bp CT tandem repeat"
exon              58872..59376
                  /gene="F21F14.20"
                  /number=2
intron            59377..59460
                  /gene="F21F14.20"
                  /number=2
exon              59461..59822
                  /gene="F21F14.20"
                  /number=3
gene              62951..64831
                  /gene="F21F14.30"
                  /complement(join(62951..63002,63099..63413,63501..63730,
63815..63927,64726..64831))
                  /gene="F21F14.30"
                  /complement(join(62951..63002,63099..63413,63501..63730,
63815..63927,64726..64831))
CDS               63815..63927,64726..64831
                  /gene="F21F14.30"
                  /codon_start=1
                  /product="ARGININE/SERINE-RICH SPLICING FACTOR RSP31"
                  /protein_id="CAB71893.1"
                  /db_xref="GI:6899884"
                  /translation="MRPVFVGNPEYETROSDBRLFDKYGVDVDMKGYAFVYFED
ERDAEDAIRLNDPFGYKRRPLSYMGAGGPRGDAKAPSNLKPKTKFVINFDP
IRKKEHDIKHFEPYKVTNVRIRRRFSPVQETQDRAKLEATQSKDIIDRVASVE
YALDDDERDRDNGKSPRSLSVTRRRPSFDYGRSPGQGRSPDYGDAASPEY
DRYKGPAYERRRSPDYGRSSDYGRSPGYDRYRSPRKTLLNAERSPFE"
exon              62951..63002
                  /gene="F21F14.30"
                  /number=1
intron            63003..63098
                  /gene="F21F14.30"
                  /number=1
exon              63413..63413
                  /complement(63099..63413)
                  /gene="F21F14.30"
                  /number=2
intron            63414..63500
                  /complement(63414..63500)
                  /gene="F21F14.30"
                  /number=2
exon              63501..63730
                  /complement(63501..63730)
                  /gene="F21F14.30"
                  /number=3
intron            63731..63814
                  /complement(63731..63814)
                  /gene="F21F14.30"
                  /number=3
exon              63815..63927
                  /complement(63815..63927)
                  /gene="F21F14.30"
                  /number=4
intron            63928..64725
                  /complement(63928..64725)
                  /gene="F21F14.30"
                  /number=4
exon              64726..64831
                  /complement(64726..64831)
                  /gene="F21F14.30"
                  /number=5
gene              65374..66567
                  /gene="F21F14.40"
                  /join(65374..65583,65679..66108,66389..66567)
                  /note="similarity to hypothetical protein - Synchocystis
sp. (Strain PCC 6803), PIR:S75899"
CDS               65374..66567
                  /codon_start=1
                  /product="putative protein"
                  /protein_id="CAB71894.1"
                  /db_xref="GI:6899885"
                  /translation="MATTHLCSTLHLPRTHPKTKINSIKRITTKSOQCKRPEISPT
PNALQIKSSSLPVALPFLPDODAAAGGRRGILGSRFAIIRHPLVWGGAAYT
LMTGYLGMWRKRVPTIOSISDIKKOLKRPUSPDGSTVADSSSPSTTHIOIGLITE
ERKRLVGSYTRDKHFDAGSYLGLGVLEAVFGVNTLYLTKGLPFGPHVAGAGITVL
WAAALAVPAMQKNDYARSILHIALNAVVLFIWQIFGLDVIKVEFTKMP"
exon              65374..66583
                  /gene="F21F14.40"

```

```

intron            65584..66567
                  /number=1
                  /gene="F21F14.40"
                  /number=1
exon              66567..66108
                  /number=1
                  /gene="F21F14.40"
                  /number=2
intron            66109..66388
                  /gene="F21F14.40"
                  /number=2
exon              66389..66567
                  /gene="F21F14.40"
                  /number=3
gene              68786..70562
                  /gene="F21F14.50"
                  /complement(join(68786..69415,69588..70562))
                  /gene="F21F14.50"
                  /complement(join(68786..69415,69588..70562))
                  /gene="F21F14.50"
                  /note="Contains Cytochrome P450 cysteine heme-iron ligand
signature AA467-476"
                  /codon_start=1
                  /product="cytochrome P450 (CYP78A9)"
                  /protein_id="CAB71895.1"
                  /db_xref="GI:6899886"
                  /translation="MATRLDPSLSLIALSKCSILTONIALSLVASTIASIALSLFF
WHPGAPAKGKFLHRRQTVITPEPRGLPVGSMISMSNTLARCIAATRAKPRAR
LWAFSIGETRVIVTNPVAKELNSVFAFADPVRESATSLMERRAIGFAFGYVMT
LKRISNHLFSPKQIKRSETQSVIANQIVCLTKQSVTKGLCFARULIKASINSM
CSVFGKEYLEBEHEBEVSELRELVEBGYDLIGTUNWDHLPWLSBPFORLRSCSNL
VYKVRFAVRILVSDHEQTRDSPDFVVLISLDGPDLSPPDIIVLWEMTEFGTDT
VAVLEMIARVILHEDIOSTYHNELDQIVGRS8AVEEDSVSLVLYTAVKEVLRILH
PGPLISMARLALITDTIIDGRVPGATPAMVMAIADHPWENPTEPERVAKE
GEVPSVYSGSDIRLAPDGGRRVCGKMLGLTYTFWATILHEFELTTSDEKTVL
SEKRLSCMANPLAKLRPRRSFV"
exon              68786..69415
                  /complement(68786..69415)
                  /gene="F21F14.50"
                  /number=1
intron            69416..69587
                  /complement(69416..69587)
                  /gene="F21F14.50"
                  /number=1
exon              69588..70562
                  /complement(69588..70562)
                  /gene="F21F14.50"
                  /number=2
gene              77018..77911
                  /gene="F21F14.60"

```

Query Match 49.8%; Score 794; DB 8; Length 142001;
Best Local Similarity 70.3%; Pred. No. 7e-220;
Matches 1238; Conservative 0; Mismatches 320; Indels 202; Gaps 5;

```

exon              199 ACCGAAAGCTATTCCTCCGCTCCAAAGGCTTCCTTGGTGAAGCATGTCACTATG 258
                  |||||
Db              70376 CAACACACCGGTATACCCGGGCCAAGAGCTTACCTTTGTGGAAGCATGTCTCTCATG 70317
                  |||||
exon              259 TCAAGACCTCTAGCTACCGAGCATGCGTATGTAGAGTGAAGTGGAGCAAGAG 318
                  |||||
Db              70316 TCAAGACCTCTAGCTACCGTTCATAGCCGACCGAGAGAAATTAGAGCGGAAGAG 70257
                  |||||
exon              319 CTCATGCTTTCAGCTTGAAGAGACTCGCGTATCGTACGTCGCAATCCGAGCATGAG 378
                  |||||

```

Db	70256	TTATGGCGTTTAACTTTGGGAGAACTCGCGTATCGTCACGTGCAATCCTGATGTAGCT	701197
QY	379	AAAGAGATTCGAAATACCCGGCTTTTGGTGCATCGACCGGTTAAAGATGGCTTACTCA	438
Db	70196	AAAGAGATTCGAACAGTCCGGTTTTCGTGACCGCCGGTTAAGAAATACGTTATTTCC	701337
QY	439	CTGATGTTTAAACAGAGCAATGTGTTTGGACAACAACGGTGTTTACTGCGACAGCTTCGC	498
Db	70136	CTCATGTTTAAACCGTCTATCGGTTTGGCTCTCTTACGCGCTTTACTGGCGAATCTTGAG	70077
QY	499	CGTATCGCTTCGAACCAATCTTTAGTACAAACAAATCGAAGACCGGAGACGACGA	558
Db	70076	AAATCCCGTCTATCATCTTTTCAGCCCGAAACAGTTTAAAGTTCCGAAGCAGAGA	700177
QY	559	CGAGTATCTCAACCCAGATGGTTGAGTTTCTTGAAAAACAAGTATGTAAGAA---CCC	615
Db	70016	AGCGTATCGCGATCAAAATCGTGAAGTGTCTCAAAACAGAGTAAACCAAAAGGCTTC	699587
QY	616	TGTTTGTGGTGGTGTGTTTAAACGGCGTCCGTTAAACAATGATGTCCTGTATTC	675
Db	69956	TGTTTCCAGTACTGATTCAAACCGCATCGCTTATATCAATGATGTCCTGTATTC	698977
QY	676	GGACACAGATATGAGCTTGAACAAACATG-----TTGAGTTACGTGAATGATC	726
Db	69896	GGAAAGAAATACGCGCTTGGAGAAAGACATGAAGAAATGAGTACCTATGATATTGGTG	698337
QY	727	GAAAGAGTTATGATTTGGTCCGAAAGGTTAATTTGATGATACCTGCTGGCTATG	786
Db	69836	GAAAGAGTTATGATTTACTCGTACACTGAAATTTGACCGCATATCTCCATGGCTCT	697777
QY	787	GAGTTGATCTTCAAGACTCCGCTCTAGATGTTCAACCTCGTACAAAGTTAAACCG	846
Db	69776	GAAATTTGATCTCAAGAAATCCGATCTAGATGCTATCTCTCCCAAAAGTTAAACCG	697177
QY	847	TTTGTATCCCGGATTTATATCCGAAACCGAATCAACCGGTTATTTGGCTGGTATTC	906
Db	69716	TTTGTGAACCGGATTTATCTCTGACACCGGAAACAACTGTATCTCACCGAGTACTTC	696577
QY	907	GTCGACGTTTGTCTTCCCTCCATGCTTCAGATTAATTTATCGACCCGACATTAATGCC	966
Db	69656	GTTGACGATATTTGCTCTCTCTCGATGCTCGAATAAATTTATCCGACCTGATTTCAATGCC	695877
QY	967	GTTCTTTG-----	974
Db	69596	GTTCTATGGGATATGTAACGTGGCTGTATTTTACTTTTGTACTGAAGAAAACTTAC	695337
QY	975	-----	974
Db	69536	AAAAATTCATATTTATTTGTACTGTATCGTTTATTAATTAATTAATTAATTAATTTTGAT	694777
QY	975	-----	974
Db	69476	AATTAATATGCAATTTTAAGTAACTAAGATTAATTTGGCAATTTGGGTTTAAATACAA	694177
QY	975	GGAGATATATTTCAAGAGAAACAACACAGTTGCGCTTATATGAGTGAATCTCGCTAG	1034
Db	69416	GGAAATATATTTCAAGAGAACTACAACGGTGTGTTTGAATGAATGAATTTCTTGCTAG	693577
QY	1035	GATGTCCTTCATTCAGATATGCAATCAACGATCAAAACGAGCTGATCAATGATGTCG	1094
Db	69356	GATGTCCTTCATTCAGATATTTCAATGACGTTTCAATATGACTTGAATCAATGTCGG	692977
QY	1095	GAAATCAAGCCCTAATGATGAATTTGACTTGGCTTCACTTCCATTTCTAACGGCTGGT	1154
Db	69296	ACGATCAAGGCGTGTCAAGAGTCTGACGCGTGTCTTATGATATTTCTAACGCTGTGAT	692377
QY	1155	GAAAGAGATTTGAGGCTTATCTCTCCAGGCCCACTTATCAATGAGGCGCTTTGGCAT	1214
Db	69236	GAAAGAGATCTTGAGGCTTACCCGACGACCACTATCTCATGAGGCCCGTTAGCAT	691777
QY	1215	AACAGACAGATCGTTATGATGTGTCTTGTTCGCGAGAGGACCAACAGCAATGATGAACAT	1274
Db	69176	CACAGACAGATATCAACGAGTGTGTGTGTTCGCGGAGGACCAACGCAATGATGAACAT	691177

OY		1275	GTTGGCCGATACGCATATCAACAAGTGGGTGAATCTTTGGATTAAACTGAG	1334
Db		69116	GTGGGCTAATTGACACCATCCAACCGTGCGGAAATCCGTTGGATTAAACCCGAC	6905
OY		1335	GTTGTGGCMAAAGAGTAGGTGAGTTTCGGTTCTTGGGTCCGATTGACACTTC	1394
Db		69056	TTTGTGAGCCAGAAGAGTAGGTGAGTTCTCGGTTCTTGGGTCTGATTGAGGCTGC	6899
OY		1395	ACCTTTCGGGTGGGTGTGTCGATTTTACCCCCCGGAGAAATCTTGTTTAACCTGAT	1454
Db		68996	ACCCTTCGGGTCGGTGTGTCGGGTTTCCCGGGAAGATCTTGTTTGAACACCGTAC	6893
OY		1455	GTTTTGACGCCAGATATGTTACATGAGTTTGAATG---GGAACCCGCGCATGTGAACG	1511
Db		68936	GTTTGGACTGCGACGCTTTTGCAATGAGTTGAATGCTACGCGTCCCATGAGAGAC	6887
OY		1512	CGTTGACTTATCTGAGAACTGAGGCTTTCTTTCGAGATGGCTAATCTCTTCTGCTAA	1571
Db		68876	CGTTGACTTGTCCAGAACTGAGGCTCTCTGTGAGATGCTAATCTCTTCTGCTAA	6881
OY		1572	ATTGCGCCGTAGGCGCAGTT	1591
Db		68816	ATTACGCCCAAGGCGCAGTT	68797
RESULT 8				
AF022463				
LOCUS		1611 bp	mRNA	linear
DEFINITION				PLN 02-MAR-1999
				Glycine max cytochrome P450 monooxygenase CYP78A3p (CYP78A3) mRNA,
VERSION		AP022463		complete cde.
KEYWORDS		AF022463.1	GI:2739007	
SOURCE				Glycine max (soybean)
ORGANISM				Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
				Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
				rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
				Glycine.
REFERENCE				1.(bases 1 to 1611)
AUTHORS				Siminszky,B., Corbin,F.T., Ward,E.R., Fleischmann,T.J. and
TITLE				Dewey,R.B.
JOURNAL				Expression of a soybean cytochrome P450 monooxygenase cDNA in yeast
MEDLINE				and tobacco enhances the metabolism of phenylurea herbicides
PUBMED				Proc. Natl. Acad. Sci. U.S.A. 96 (4), 1750-1755 (1999)
				99145622
				9990096
REFERENCE				2.(bases 1 to 1611)
AUTHORS				Siminszky,B., Dewey,R.E. and Corbin,F.T.
TITLE				Direct Submission
JOURNAL				Submitted (04-SBP-1997) Crop Science, North Carolina State
				University, Box 7620, Raleigh, NC 27695, USA
FEATURES				Location/Qualifiers
source				1..1611
				/organism="Glycine max"
				/mol_type="mRNA"
				/db_xref="taxon:3847"
				1..1611
				/gene="CYP78A3"
				20..1591
				/gene="CYP78A3"
				/note="cytochrome P450 monooxygenase"
				/codon_start=1
				/product="CYP78A3p"
				/protein_id="AB94592.1"
				/db_xref="GI:2739008"
				/translation="MTSHDDNMITALTSTKCTQENLAWYLIMSLMILMPYYMSH
				PGCPANGKYVTPPLISLTIDGPKGPILTSMGIMSTLAHRRLDAATAORAKRMPS
				LCHTRIVTCHPVAKEILINSVFADRPKBKSYSLMFRALIGPAHYGVMBEIRIA
				SNLFPCROTKASELORSQAQMWHIILNKRRSRLLRVRAIKKASLSMMSCSVFGE
				VRIADHPNSMEDDGIILVDGGYDLGLFNNAADIPLIAHFDAQIRKCSLVPMVRF
				VGTLIEHRSKTEINRDPDYVLISLPEDQLSDSMIALVLMEMIERGIDTVAVILEW
gene				
CDS				

Tue Jan 20 17:43:21 2004

us-10-022-025a-2.rge

Page 14

ILAMALHPHVSQVSELDNAVGRKARAVAEEDVAVMTVLPAVKEVILRHPPELIS
WARSINDITIDGTHVPHAGTTAMVNTAICRDPHVMDPLEPMPERBTRVGAERSI
LGSDFRLAPFSGRRACGKTLGMATVNFVASHLEFEWPSDEKVDULEVLKASS
EMANPLTVKXPRRG"
BASE COUNT 369 a 457 c 407 g 378 t
ORIGIN

Query Match 40.9%; Score 651.6; DB 8; Length 1611;
Best Local Similarity 66.5%; Pred. No. 1.1e-186;
Matches 1025; Conservative 0; Mismatches 484; Indels 33; Gaps 5;

QY 61 AGCCAAACCAACCTTCTCCTCCTCCGCTGCAATCATCTGCTGCCATATCT 120
DB ACCCAAGAAACCTTGCATGGGTCTTTGATCATGGGCTCACTGGTTAACATGACT 133
QY 121 CTCTTCTTATGACCTATCCGGTGGAGCTGTTGGGGGAAATACCTTGGCGGTTA 180
DB 134 TTCTATTACTGGTCAACCCCGGTGGCTGCTGGGGAAGTACTACAC-----TA 186
QY 181 ATATCCGCTTCATACAAACCGAAACGTTATCCCGGTCCAAAGGCTTCCCTTGGTT 240
DB 187 CTCTCCCCCTTTCAA-----TCATTCGGGTCCCAAGGCTTCCCTCTTATT 235
QY 241 GGAAGCATCTACTCATGTCAAGCACTAGCTCACGACGAATGCTGATGCACTGAG 300
DB 236 GGAAGCATGGGCTCATGACTTCC--CTGGCCCATCACCTATTCGACGCGCGCC 292
QY 301 AAATTCGAGCAAGAGGCTCATGGCTTCACTTAGAGAGACTCGGCTGATGCTACG 360
DB 293 ACATGACAGCCAAAGGCTCATGGCTTTAGTCTCGGCAACACGCTGATGCTACG 352
QY 361 TCGAATCCCGACGTAGCGAAGAGATTCGAATAGCCCGTTTCTGATGACCGGTT 420
DB 353 TCGAATCCCGACGTAGCGAAGAGATTCGAATAGCCCGTTTCTGATGACCGGTT 412
QY 421 AAAGATTCGGCTTACTCATGTGATGATGATGATGATGATGATGATGATGATGAT 480
DB 413 AAAGATTCGGCTTACTCATGTGATGATGATGATGATGATGATGATGATGATGAT 472
QY 481 TACTGGCGAAGCTTGGCGGATGCTGATGATGATGATGATGATGATGATGATGAT 540
DB 473 TACTGGCGAAGCTTGGCGGATGCTGATGATGATGATGATGATGATGATGATGAT 532
QY 541 AGAGCCGAGACGAGACGAGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 533 GCCTGAGACTCCAAAGCTCTCAATAGCGCGCCCAATAGTTCATCTCT--AAATAC 589
QY 601 AGTAGTACGAAACCTGTTTGTGCTGAGTTGCTTAAACGCGCTGCTTAAACATG 660
DB 590 AACCGCAACCGAGCTTAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 649
QY 661 ATGTGCTCTGATTCGACAGAGTATGAGTGTGAAAAAACA-----TGTGAG 711
DB 650 ATGTGCTCTGATTCGACAGAGTATGAGTGTGAAAAAACA-----TGTGAG 709
QY 712 TTAGTGAATGATGACAGAGGATGATGATGATGATGATGATGATGATGATGATG 771
DB 710 CTGGAATATATGATGACAGAGGATGATGATGATGATGATGATGATGATGATGAT 769
QY 772 CTTCCTTGGCTATGAGATGATGATGATGATGATGATGATGATGATGATGATG 831
DB 770 CTTCCTTGGCTATGAGATGATGATGATGATGATGATGATGATGATGATGATG 829
QY 832 CCAAGGTAAACCGGTTGATCCCGATATATCCGAACACCGATATCAACCGGTGAT 891
DB 830 CCATGATGAAACGCTTTCGTCGACAAATATGCTGACACCGACCTGTAAACCGAA 889
QY 892 TTGCTGCTGATTCGTCAGCTTTGCTCTCCCTCAGTGTGATGATGATGATGATG 951
DB 890 ACCAATGCTGATTTGTTGACGCTTCTCTCTCCCGAACCCTATCATATATGAC 949
QY 952 CCGGACATATCGCGCTTCTTGGGAGATGATGATGATGATGATGATGATGATGATG 1011

DB 950 TCCGATCATGCTGCTACTTGGGAAATGATATTCAGAGAAACGACACCGTACG 1009
QY 1012 TTAATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1071
DB 1010 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1069
QY 1072 AACGAGCTGATCAAGTATGCTGGGAATCAAGAGCCCTTATGATGATGATGATG 1131
DB 1070 GAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1129
QY 1132 CTTCCTATCTAAGCGCTGATGATGATGATGATGATGATGATGATGATGATG 1191
DB 1130 ATGACGATCTTACAGCGGTGATGATGATGATGATGATGATGATGATGATGAT 1189
QY 1192 CTATCATGAGCGCTTGGCCATTAACAGACAGATGATGATGATGATGATGATG 1251
DB 1190 CTATCATGAGCGCTTGGCCATTAACAGACAGATGATGATGATGATGATGATG 1249
QY 1252 GGAACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1311
DB 1250 GGAACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1309
QY 1312 CTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1371
DB 1310 CCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1369
QY 1372 CTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1431
DB 1370 CTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1429
QY 1432 ATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1491
DB 1430 ACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1489
QY 1492 GGAACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1551
DB 1490 GGAACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1549
QY 1552 GCTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1593
DB 1550 GCTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1591

RESULT 9
ATH552847/c
LOCUS
DEFINITION
Arabidopsis thaliana T-DNA flanking sequence, left border, clone 345B11.
ACCESSION
AJ552847
VERSION
AJ552847.1
KEYWORDS
left border; T-DNA flanking sequence.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; rosoids; eurosid II; Brassicales; Brassicaceae; Arabidopsids.
1
Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Craud, C., Derose, R., Pelletier, G., Lepoint, L., Caboche, M., and Lecharny, A.
T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites
EMBO Rep. 3 (12), 1152-1157 (2002)
JOURNAL
MEDLINE
22363535
PUBMED
12446565
REFERENCE
2 (bases 1 to 740)
Balzerque, S.
Title
Direct Submission
Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE
JOURNAL
AUTHORS
COMMENT
The PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment (b) resulting from the PCR were directly sequenced from the left or the right border

Db 84386 TTATATCTCTCATTACACCGGAAATTCATTCCTCCGGCTCTAGAGGTTT 84327
 QY 232 CCTTTGGTGAAGAGATGTCATCTCATGTCAAGCACTCTAGCTCAACCGAATTCGTAT 291
 Db 84326 CCACTTGTGGAGAGATGAGCTTAGGTCAAGCCAGTGGCTCATAGGCGAATAGGGTCT 84267
 QY 292 GCAGCTGAGAAATTCGAGAGCAAGAGGCTCATGGCTTTCAAGCTTAGAAGAGCTGGCGT 351
 Db 84266 GTGGCTGAGATGATTAACCGCCAGCGGCTCATGGCTTTAGCTCGGTATCTAGAGGTG 84207
 QY 352 ATGTGACCTGCAATCCCGAGTGAAGAAAGATCTGAAATAGCCGGCTTTTGTGTAT 411
 Db 84206 GTGGTACGTGTCTCTCCGTGGCAAGAGATCTAAACAGTTCCGCTTTTGTGTAC 84147
 QY 412 CGACCGGTTAAAGATTCGGCTTACTCACTGATGTTTAAACAGCAATGGTTTTGAACA 471
 Db 84146 CGACCGGTTAGCAAAACCGCTTACGGTTGATGTTTAAACGAGCCATGGGATTTGCTCC 84087
 QY 472 CACGGTCTTACTGGCGAAGCTTCCGCTTACGCTTCCGACCATCTCTTTAGTACAA 531
 Db 84086 AATGTACTTATTTGGCGTACGCTGCGGTAGGCTCGAACCATCTTTTAAACCGAAG 84027
 QY 532 CAATTCAGAAAGCCGAGACGCAACGAGATCTCAAGCCAGATGGTTGAGTTCTT 591
 Db 84026 CAATTCAGATCGAGAGATCGAGACGAGTGTATGCGACTCAGATGGTGAATGCGTT 83967
 QY 592 GAAAAACAGAGTACTTAAACAACTCTGTTTGTCTGATGTTGCTTAAACAGGCGTGGCTT 651
 Db 83966 G---CAGTAAACCTTAATCCGCGTGCAGTGGTATTTGCTCAAAACAGGCGTGGT 83910
 QY 652 AACAACAGATGCTCTGATTTGCAAGAGATGAGCTTGAACCA----- 704
 Db 83909 TGTAACTGATGGTGTGTTTGTGGAGAGAGATGATTTGAGTCAAAATTAACAATTG 82850
 QY 705 -----TGTGATTAAGTGAATGTCGAAGAGATTAATTTGCTCGAAGCTTGAAT 759
 Db 83849 GAATCTGAATGCTTAAAGGGTGTGTTGAGAGAGGATCAATCTTCAAGTACGTTAAAT 83790
 QY 760 TGAAGTATCACTCTCTGCTGCTATCGAGATTTGATCOTCAAAAGCTCCGATCTAGAT 819
 Db 83789 TGGACCGACATCTCTCTGCTTACCGGTTTAAATTTCAACAAATCCGCTTTAGTGC 83730
 QY 820 TCCACACTCTGATCCAAAGATTAACCGGTTGATCCGAGTATATCCGAACACCGTAT 879
 Db 83729 TCGACAGCTCGTACCGAAAGTAACTGTATTTAGACCGTATCATATGAACAACGCTCT 83670
 QY 880 CAACACCGTATTTGCTCGTGTATTTGCTGACAGTTTGTCTCCCTCCCATGGTTCAGAT 939
 Db 83669 GCCACGGGTAACTTCT-----TTACAGTAACTTCTTCTCAAGTTCCGAA 83622
 QY 940 AATATATCCGACCCGAGATATCGCCGCTCTTTG----- 974
 Db 83621 AATATATGAGATCCGACATGTTGCTTCTTTGGTGAAGTTTGTCTTTTAACTACTA 83562
 QY 975 ----- 974
 Db 83561 TTCAATCCGATATCATGATATCAACATTAAGTCACTAGATTAACATTAAGAAATCC 83502
 QY 975 ----- 974
 Db 83501 AATAAATTTACTAATTAAGTATAGTATTAAGATTAACAATGATTAACATTTGTTGTT 83442
 QY 975 ----- 974
 Db 83441 TTATATCAATCATGAGAAATGATTAAGAAATTTCTCTTACGATATTAACCTTTTGGT 83382
 QY 975 -----GGAGATGATATTCAGAGGAACAGACAGTTCGGCTTTAAATCGAGTGAATCC 1027
 Db 83381 CCGTTTAGGAATGATTTTATAGGGAACGAGACATGTTGCGGTTTGGTGGAGTGGTGC 83322
 QY 1028 TCGCTAGATGGTCTCTTCAATGATATGCAATCAACGATTAACAAACGAGCTGATCAAG 1087
 Db 83321 TAGCGAGATGATGATGATCCCAAGTTCAATTAACGCTTCCACATGAGCTTGAACGAG 83262

QY 1088 TAGTCGGAAATCAAGAGCCCTTAAATGATCTGACTTGGTCACTTCAATCTAAGCG 1147
 Db 83261 TCGTTGGAGATCAAGAACCGTGGATGATGAGCTTCCATCACTACATGATTAACGG 83202
 QY 1148 CTGTGTGAAGAAGATATTGAGGCTTCACTCCCGAGCCCACTTCAATCAATGGCCCGTT 1207
 Db 83201 CTATGATCAAGAAGATGTTAGGCTGCATCAACAGGATCACTGCTTTTGGGACAGAC 83142
 QY 1208 TGCCCATTAACAACAAGATCTTATAGTGTCTTTGTCGGGACGGAACAAGCAATGG 1267
 Db 83141 TGTCTATTAACAGACACTTCCCTAATGATATACAGTGGCGGCTGGAGACACCGCATGG 83082
 QY 1268 TGAACATGAGGCGGATTCGATATCGATATCCACAGTGGGTTGATCTTTGAGTTTAAAC 1327
 Db 83081 TCACATGAGGCTATAGACAGTACCCACACAGTGGAGAGATCTTTAGATTTAAGC 83022
 QY 1328 CTGAGAGTTCGTGGCAAAAGAGGTGAGTGTGAGTTTCCGTTCTTGGTGGATTTGA 1387
 Db 83021 CTGAGAGTTCGTGGCTAAAGAGGCTGAGTGTCTGTTTCCGGTCCGATCTGA 82962
 QY 1388 GACTTGCACCTTTGGGCGGTGGGTGGATTTGCCCCGGAGAAATCTGTTTACTA 1447
 Db 82961 GGTGGCACCGTTGCGGCTGAGGTTAAGAGGTTTCCCTCGAAAGATTTGGACCTTACA 82902
 QY 1448 CCGTATGTTTGGACGCGCATGATGTTTACATGAGTTGAATGGGAGAC--GTCCGATG 1504
 Db 82901 CGGTGTGTTTGGGTTGCAACGCTCTTGCATGATTTGAGTGGCTTCCAGCGTGAAG 82842
 QY 1505 GTTACGGGCTGACTTATCTGAGAACTGAGAGCTTTCTTGGAGATGCTATCTCTTC 1564
 Db 82841 CTAAACCTCCAGATCTCTCGAGGTTTGAAGGCTCTCGTGTGATGAGTCTGTCCACTCA 82782
 QY 1565 CTGCTAAATTCGCGCGGTAGCGCA 1588
 Db 82781 TCGTTAACGTAAGCTCAAGGGCTA 82758

RESULT 11
 AC023628/c
 LOCUS
 DEFINITION
 Arabidopsis thaliana chromosome I BAC FEF3 genomic sequence,
 complete sequence.
 AC023628
 VERSION
 AC023628.3 GI:7212005
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 104001)
 2 (bases 1 to 104001)
 3 (bases 1 to 104001)
 4 (bases 1 to 104001)
 5 (bases 1 to 104001)
 6 (bases 1 to 104001)
 7 (bases 1 to 104001)
 8 (bases 1 to 104001)
 9 (bases 1 to 104001)
 10 (bases 1 to 104001)
 11 (bases 1 to 104001)
 12 (bases 1 to 104001)
 13 (bases 1 to 104001)
 14 (bases 1 to 104001)
 15 (bases 1 to 104001)
 16 (bases 1 to 104001)
 17 (bases 1 to 104001)
 18 (bases 1 to 104001)
 19 (bases 1 to 104001)
 20 (bases 1 to 104001)
 21 (bases 1 to 104001)
 22 (bases 1 to 104001)
 23 (bases 1 to 104001)
 24 (bases 1 to 104001)
 25 (bases 1 to 104001)
 26 (bases 1 to 104001)
 27 (bases 1 to 104001)
 28 (bases 1 to 104001)
 29 (bases 1 to 104001)
 30 (bases 1 to 104001)
 31 (bases 1 to 104001)
 32 (bases 1 to 104001)
 33 (bases 1 to 104001)
 34 (bases 1 to 104001)
 35 (bases 1 to 104001)
 36 (bases 1 to 104001)
 37 (bases 1 to 104001)
 38 (bases 1 to 104001)
 39 (bases 1 to 104001)
 40 (bases 1 to 104001)
 41 (bases 1 to 104001)
 42 (bases 1 to 104001)
 43 (bases 1 to 104001)
 44 (bases 1 to 104001)
 45 (bases 1 to 104001)
 46 (bases 1 to 104001)
 47 (bases 1 to 104001)
 48 (bases 1 to 104001)
 49 (bases 1 to 104001)
 50 (bases 1 to 104001)
 51 (bases 1 to 104001)
 52 (bases 1 to 104001)
 53 (bases 1 to 104001)
 54 (bases 1 to 104001)
 55 (bases 1 to 104001)
 56 (bases 1 to 104001)
 57 (bases 1 to 104001)
 58 (bases 1 to 104001)
 59 (bases 1 to 104001)
 60 (bases 1 to 104001)
 61 (bases 1 to 104001)
 62 (bases 1 to 104001)
 63 (bases 1 to 104001)
 64 (bases 1 to 104001)
 65 (bases 1 to 104001)
 66 (bases 1 to 104001)
 67 (bases 1 to 104001)
 68 (bases 1 to 104001)
 69 (bases 1 to 104001)
 70 (bases 1 to 104001)
 71 (bases 1 to 104001)
 72 (bases 1 to 104001)
 73 (bases 1 to 104001)
 74 (bases 1 to 104001)
 75 (bases 1 to 104001)
 76 (bases 1 to 104001)
 77 (bases 1 to 104001)
 78 (bases 1 to 104001)
 79 (bases 1 to 104001)
 80 (bases 1 to 104001)
 81 (bases 1 to 104001)
 82 (bases 1 to 104001)
 83 (bases 1 to 104001)
 84 (bases 1 to 104001)
 85 (bases 1 to 104001)
 86 (bases 1 to 104001)
 87 (bases 1 to 104001)
 88 (bases 1 to 104001)
 89 (bases 1 to 104001)
 90 (bases 1 to 104001)
 91 (bases 1 to 104001)
 92 (bases 1 to 104001)
 93 (bases 1 to 104001)
 94 (bases 1 to 104001)
 95 (bases 1 to 104001)
 96 (bases 1 to 104001)
 97 (bases 1 to 104001)
 98 (bases 1 to 104001)
 99 (bases 1 to 104001)
 100 (bases 1 to 104001)

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Submitted (16-FEB-2000) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 3 (bases 1 to 104001)

ALRSVTPNISKDKKALUGDANSYINELHAKUKMEAEERBRIGSNPISIDSDIN
 QTSGEDVTVRINCPLSHPSARI FFAFESKVEVINSNLESDVTLMTFFVKSBLT
 KEXLISALSRBQNSVOSRUSGR
 53198..54813
 /gene="F6P3.8"
 /join(53198..54103,54187..54813)
 /note="highly similar to dbj|BAA2894.1, cytochrome P450
 from Petunia x hybrida"
 /codon_start=1
 /product="putative cytochrome P450"
 /protein_id="AAF9733.1"
 /db_xref="gi:966139"
 /translation="MTILVASLFAVILINVLIRMLKASAKRAQLRPPRPRLILGN
 LLOGLPRLHDLASLCKRGPPLYLRLGNDATITTDPRIRLILRODVSRPKT
 LAAYHLAYGGDVALAPMGPBMRKMRICHEHLITRLESFTTORAEPRYLIRDP
 KRSETGRINLKEVLGAFSMANVTRMLGQFPGPSLVSPEKAEFLHITKLFML
 GYLIGDYLPEFRWVDPSCGEKRDVEKRVDFHTKIDENRALLEDKNDKMDP
 VDLISLPGENKAMEDVHILKLIQDMLAADTSVTEMMAEAIQPRMYKIQ
 EELDNVGSNRMVDESLVHLNLCVRETFPMHAPGPELIPHSVRAATTINGYIP
 AKRVFINTHGLGNTKIMDVEDFPERHMPVSGSRVISHGPDFKILPFAGKRX

Query Match 29.8%; Score 474.4; DB 8; Length 104001;
 Best Local Similarity 59.6%; Pred. No. 1,7e-132;
 Matches 1087; Conservative 0; Mismatches 486; Indels 251; Gaps 7;

QY 1 ATGGCTACGAACCTGGAAGCTCCTTAATCTTTCCTTTGTCCTCAATGAGGCTCTA 60
 Db 25772 ATGAGAACCGGAATGAAAGTTTGTGGTATTTCCTTCCTGCAATAATCAATTTAC 25713
 QY 61 AGCCAAACCAACCTTGCCTTCTCCTCTCGCGCTCAACATCATGCTGCGCAATCT 120
 Db 25712 ATGCAACAACTTTTGTCTTCTCCTTCTCGTCCCATCTCATCTGCTTACCATTAAC 25653
 QY 121 CTCTTTTATGACCTATCCCGGTGGAAGCTTGGGGGAAATACCTCTTC--GGCGG 177
 Db 25652 ATCGTATTTGTCTACTCCGAGTGAACGCTGGGGAAATATCTTCTTCACTGCGCG 25593
 QY 178 TTAATATCCGCTTCATAC-----AAACCGGAAACGTTATCCCGTCCAAAGGCTTC 231
 Db 25592 TTATATTTCTCGATTAACAACGAATATCAATCTCATCCCGGCTTAAGAGGCTTT 25533
 QY 232 CTTTGTGTTGAAGATGTCATCTATGCAAGCACTAGTCAACCAAGAAATGGCTGAT 291
 Db 25532 CCACTTGTGGAAAGATAGGCTTATGTCAGCAACGTCGCTCATACGCGATAGCGCT 25473
 QY 292 GCAGCTGAGAAATTCGAGCCAAAGAGCTCATGGCTTCACTTATAGAGAGACTCGGCTG 351
 Db 25472 GTGGGTGAGATGAGTAAGCCAAAGGCGCTCATGGCTTATAGCTCGGATATCAAGGTG 25413
 QY 352 ATGCTACGTCAATCCGACGTAGCCAAAGAGATTCGAATAGCCGCTTTTGTGCTGAT 411
 Db 25412 GTGTGAGAGTGTCACTGCGCTGCAAGAGAGATCACTAACAGTTCGTTTTTGTGCAC 25353
 QY 412 GCACCGGTTAAAGATCGGCTTACTCACTGATGTTTAAACAAGCAATGTTTGGACCA 471
 Db 25352 GCACCGGTTGACGAACCGCTTACGCTTATGATGTTTAAACGACCAATGGGATTTGCTCC 25293
 QY 472 CACGGTGTTTACTGCGCAACGCTTTCGCGTATCGCTTCGAMCCATCTCTTATAGTCAAAA 531
 Db 25292 AATGTAATCTTATTTGCGTACGCTGCGTGGGTATAGCTCGAACCAATCTTTTAAACCGGAG 25233
 QY 532 CAATTCAGAGAGCCGAGACGCAAGAGAGATCTCAAGCCAGATGTTGAGTTCTT 591
 Db 25232 CAATTCAGATTCGAGAGATGAGACGCGGTGATAGGACTCAATGATGTAATCCGTTT 25173
 QY 592 GAAAAACAGAGTATGTAACGAACCTGTTTGTTCGTGAGTGTCTTAAACGCGCTGCTT 651
 Db 25172 G---CACGTAAACCTTAATCCGCGTGTGCAAGTGCATTTGCTCAAAAACGCGCTG 25116
 QY 652 AACCAATGATGTGCTCTTATTTGCAACAAGATGAGCTTGAAGAAAAACA-----704
 Db 25115 TGTACATATGAGGTGTGTTTTGCGAGAGAGATATGAATGAGTCAATTAACAACCTTG 25056

QY 705 -----TGTAGTTACGTAAGAAATGGTGAAGAAAGTTATGATTTGCTCGAACCTGTAAT 759
 Db 25055 GAATCTGAATGCTTAAAGGTTTGTGTTGAGAAAGGATGCAATCTTCAAGTACCTTAAT 24996
 QY 760 TGAAGTATCACTTCTCTTGGCTATCGGAGTTGATCCTCAAGACTCCGCTTGAATGT 819
 Db 24995 TGAACCAACATCTTCTCTTGGTTAGCCGGTTAATATTCACAAATCCGGTTTAGGTGC 24936
 QY 820 TCACACTCTGTAACCAAGTAAACCGTTGTATCCCGGATTTATATCCGACACCGTAAT 879
 Db 24935 TCGAGCTTCGTACCAAGGTAATCTGTTATGAGCCGCTATCATATGAAACACGTCCT 24876
 QY 880 CAACCGGTGATTTGCTCGTATTTCTGACAGCTTTGCTCTCCATGATGTTCAAGT 939
 Db 24875 GGCACGGTAATCTTCT-----TGACATGTTACTTCTTCAAGGTTCAAGAA 24828
 QY 940 AATTTATCCGACCCGAGATTAATGCCCTTCTTTG----- 974
 Db 24827 AATTTATCAGATTCGACATGAGTGTCTGCTTGGGTAAGTTTGTCTTTAATCACTA 24768
 QY 975 ----- 974
 Db 24767 TTCAATCCGTATATGATGATATCAACATTAGTCAACTAGATTAACATAAGAAATCC 24708
 QY 975 ----- 974
 Db 24707 ATAAATTTACTAATTAAGTAAGTATTAAGTATTAACAATGATTACAATGTTGGTT 24648
 QY 975 ----- 974
 Db 24647 TTATATCAATATGAGAAATGATTAAGAAATTTCTCTTACGATTAATCTTTTGGT 24588
 QY 975 -----GGAGATGATATTCAGAGAAACAGACACAGTTGCGGTCTTAAATCGATGATCC 1027
 Db 24587 CGGTTAGAAATGATATTTAGGGAACGACACAGTTGCGGTTTGGTGTGCGTGGTGC 24528
 QY 1028 TCGTGTGATGTGCTCTTATCCAGATNAGCATCAACGGTACAAACGAGCTGATCAAG 1087
 Db 24527 TAGCGAGATGTGATGATCCCAAGATTCATTAACGATCAACATAGCTTGAACGAG 24468
 QY 1088 TAGTCGGAATTCAGAGCCCTAGATGATGATCTGACTTGTCTTCACTTCCATATCTAACG 1147
 Db 24467 TCGTTGGAGATCAAGAAACGATGATGATGATCACTTCCATCACTCACTGATCTAACG 24408
 QY 1148 CTGTGTGAAAGATTTGAGGCTTCACTCCAGGCCCACTTCTATCATAGGGCCGCTT 1207
 Db 24407 CTATGATCAAGAAAGTGTGAGGCTGATCCACAGCTCACTTCTTGTGGACAGAC 24348
 QY 1208 TGGCATTAACAGACAGATGTTGATGTCTGCTTGTTCGCGACGAGACCAACAGCAATG 1267
 Db 24347 TGTCTTAACAGACATTCGATGATGATGATATCACTGCGGCTGGGACCAACGCGATG 24288
 QY 1268 TGAACATGTGGCCGATATGATGATATCCACAGTGTGGTATCCTTTGAGATTAAAC 1327
 Db 24287 TCAACATGTGGGCTATAGACGATGACCAAGTGTGGAGAGATCTTTAGATTTAAAC 24228
 QY 1328 CTGAGAGGTTGCTGCAAAAGAGATGAGTGAAGTTTTCGTTCTTGTGGTGGATTTGA 1387
 Db 24227 CTGAGAGGTTGCTGCTTAAGAGGTTGAGAGCTGATGATCTTCTGTTTTCGCGGTG 24168
 QY 1388 GACTTGACCTTTCGCGTGGGTGCTGCGATTTTGGCCCGGAGAAATCTTGTGTTACTA 1447
 Db 24167 GATTGCAACGCTTCGCGGTGGGTGAGAGGATTTGCGCTGAAAGATTTTGGACTTAACA 24108
 QY 1448 CCGTTATGTTTGGAGCGGATGATGATCATGATGATTAATGAGGAGAC---GTCCGATG 1504
 Db 24107 CCGTGTGTTTGGGTGCAACGCTCTGATGATGATTTAGATGCTTCCATAGGTGAGAG 24048
 QY 1505 GTAACGCGTGAATTAATCTGAGAAATGAGGCTTCTTTCGAGATGAGTAAATCTCTTC 1564
 Db 24047 CTAAACCTTCAGATCTCTCGAGAGTTTTGGGCTCTGCTGTGATGATGCTTGTCCACTCA 23988

QY 1565 CTGCTAATTCGGCCGTAGGCGCA 1588
 Db 23987 TCGTTACGTACGTCAAGCGCTA 23964

RESULT 12
 LOCUS AR070567 1799 bp DNA linear PAT 18-FEB-2000
 DEFINITION Sequence 3 from patent US 5907082.
 ACCESSION AR070567
 VERSION AR070567.1 GI:7221455
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1799)
 AUTHORS O'Neill, S. and Nadeau, J.
 TITLE Ovine-specific gene expression
 JOURNAL Patent: US 5907082-A 3/25-MAY-1999;
 FEATURES
 source 1..1799
 /organism="unknown"

BASE COUNT 456 a 362 c 428 g 548 t 5 others

Query Match 22.3%; Score 355; DB 6; Length 1799;
 Best Local Similarity 58.0%; Pred. No. 2.8e-96;
 Matches 744; Conservative 0; Mismatches 500; Indels 39; Gaps 5;

QY 304 TTCGAGCCAAAGAGCTCATGCTTCACCTTAGAGAGACTCGCGTATCGTCACGTGC 363
 Db 273 TTCGAGCCCTCCGCTCATGCTTCCTGCTGCGCTCCTCATCGCTTCACTGCTTCAAGC 332
 QY 364 AATCCGACGTAGCGGAAGATCTGAATAGCCGGCTTTTGTGATCGACCGGTTAA 423
 Db 333 CACCCGAAACCGCAAAAGATCTCTTCAAGCCGCTTCGTGATCGGCCCATTTAA 392
 QY 424 GAATGCGCTTACTCATGTAGTTTAAAGAGCAATGTTTGGACCAAGGTTTAC 483
 Db 393 GAATCAGCATGCAATCTTCTGTTAATCGCGCTTAGGGTTTCCCAATTTGGGATTC 452
 QY 484 TGGCGAACCTTCGCGGTATCGCTTGAACCATCTCTTAGTAAACAAATCAGAGA 543
 Db 453 TGGAGAACTCGAGAGATTCGTCACATATCTTTCACTCGCGGAGATTTCATCG 512
 QY 544 GCGGAGCGCAAGAGAGATCTCAAGCAAGATGTTGAGTTCTTGAACAGAGA-- 601
 Db 513 TTTCAGAGACGAGAGATGATGTCGAGAGATGTCGGGATTTGAAGAAATGATG 572
 QY 602 -GTAGTACGAAACCTGTTTGTTCGTGATGTCCTTAAACGCGCTGCTTAAACATG 660
 Db 573 GAGAGAAATGAGATGTTAGAGTGAAGAGAGATGTTGCACTACGGGCTTTGAATTAATC 632
 QY 661 ATGTGCTCTGATTCGAGCAAGATGAGCTTGAACAAACCATTTAGTTAGTGA 720
 Db 633 ATGTGATCTGTTTGGGAAAGTTGATTTGCAAGAGATGAGGGGTTGAGCTGAG 692
 QY 721 ATGATC--GAGAGAGATGATGATTTGCTCGAAGCTTGAATGATGATCACTTCT 777
 Db 693 TTGATCCTTAAGAGAGATGATGATTTGATGAGATCTTCACTCGGGGATCATTTGCT 752
 QY 778 TGCTATCGAGATTTGATCTCAAAGATCCGCTTAGATGTTTCCACTCTGTAACAG 837
 Db 753 CTTTGGGATGTTAATTTGCAAGGTGAAGAGAGATGCAACACTTGTGGCTAAG 812
 QY 838 GTAACCGGTTTGTATCCGATTTATATCCGAACACCGTAATCAAC----- 885
 Db 813 GTCAATGATTTGTGAAGAGATCATGACGACATTAAGAGAGACCAACCGCGTAAGG 872
 QY 886 -----GATATTTGCTGATTTGCTGACGTTTGTCTCTCCCTCAATGTTCAAT 939
 Db 873 ATTGATGAGGTGAAGGTGAAGATTTGTTGATGATGCTCTTGGTTGAGAGAGAAAT 932

QY 940 AATATCCGACCCGACATTAATCGCGTTCTTTGGAGATGATATTCAGAGAACAGAC 999
 Db 933 AGACTCTCAGAACTATGATATGCTCGAGTTCTTTGGAAATGATCTTTAGAGAACTGAT 992

QY 1000 ACACTGCGGCTTTAATCGAGTGAATCCCTGCTAGAGATGCTCTTATCCAGATGCA 1059
 Db 993 ACTGTTGCCATCTTATGATGATGAGATGAGCTTGGATGATGATGATCTTATCTTCA 1052

QY 1060 TCAACGTACAAACAGCTGATCAAGTATGTCGGAATTCAGAGCCCTAGATGATCT 1119
 Db 1053 TCGAAGGCAACAATTGATGATTTCTGTCGTTGACTCTTCAAGGCAAGATTTGATCT 1112

QY 1120 GACTGCTTCACCTTCATATCTAAGCGCTGTGTAAGAAAGATATTAGGCTTCACTCT 1179
 Db 1113 GATATCCAGACTCTTATCTTATCTCAATCTATAGTAAAGAAACCTTGAATGATCTCT 1172

QY 1180 CCAAGCCACTTCTATCATAGGCGCGTTTGGCCATTAACAGACAGATCGTTAGTGTCT 1239
 Db 1173 CTTGGGCTTATTTGTAAGGCTCGCTAGCTATCATGACGCTTCTGATGATGCTAC 1232

QY 1240 CTGTTCCGCGCAGGAGCCACAGCAATGTGAACATGTGCGCGCTATCGATATCCAC 1299
 Db 1233 ATGATTCCTGCTGGAGCACTGCAATGTGAACATGTGGGCAATACATGACGATGCTC 1292

QY 1300 GTGTGGTGTATCTTGTAGATTAACTGAGAGGTTGTGCGCAAAAGAGTGAAGTG 1359
 Db 1293 AACTGGCTAGGCTTAACAAATTAATCTGATGATTCATCATGATGAAGT----- 1344

QY 1360 GAGTTTCGTTCTTGGTGGATTTGAGACTTGCACCTTCGGGTGCGGTCTGGAAT 1419
 Db 1345 ----TCAATATCTTGGTTCGATTAAGGTTGACCTTGTGCTCCGTAAGAGAT 1400

QY 1420 TGCCCCGGAAGATCTGTTTACTACCTGATATGTTTGAAGCGCGATGATTAAT 1479
 Db 1401 TGCCCTGCAAAACGATGAGATGCTGCGACGTTATCTTGTGCTCAGTGTCTGAA 1460

QY 1480 GAGTTTAAATGGGACCGCTCGATGTTAAGCGGTTGACTTATCTGAACAACTGAGCTT 1539
 Db 1461 AGCTTCAATTTGCTCTTC--GAGAAATGATGATGATTTGCTGATGCTTAAAGATG 1517

QY 1540 TCTTGGAGATGCTAATCTCTCT 1562
 Db 1518 TCTCTCGAGATGAAGATCTTT 1540

RESULT 13
 PSU34744 1799 bp mRNA linear PLN 02-DEC-1996
 LOCUS PSU34744
 DEFINITION Phalaenopsis sp. 'hybrid SM9108' cytochrome P-450 (CYP8A2) mRNA,
 complete cds.
 ACCESSION U34744
 VERSION U34744.1 GI:1173623
 KEYWORDS
 SOURCE

ORGANISM Phalaenopsis sp. SM9108
 Phalaenopsis sp. SM9108
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophytes;
 Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
 Epidendroideae; higher Epidendroideae; Vandeeae; Aeridinae;
 Phalaenopsis.
 REFERENCE 1 (bases 1 to 1799)
 AUTHORS Nadeau, J. A., Zhang, X. S., Li, J. and O'Neill, S. D.
 TITLE Ovine development: identification of stage-specific and
 tissue-specific cDNAs
 JOURNAL Plant Cell 8 (2), 213-239 (1996)
 MEDLINE 96351416
 PUBMED 8742709
 REFERENCE 2 (bases 1 to 1799)
 AUTHORS Nadeau, J. A.
 TITLE Direct Submission
 JOURNAL Submitted (23-AUG-1995) Jeanette A. Nadeau, Section of Plant
 Biology, University of California, Davis, Davis, CA 95616, USA
 FEATURES
 source 1..1799

```

/organism="Phalaenopsis sp. SM9108"
/mol_type="mRNA"
/strain="hybrid SM9108"
/db_xref="taxon:54872"
/clone="040"
/tissue_type="pollen tubes"
/clone_lib="megaportocyte-stage ovule lambda Zap11
library"
1.1799
/gene="CYP78A2"
291.1571
/gene="CYP78A2"
/codon_start=1
/product="cytochrome P-450"
/protein_id="AA87221.1"
/db_xref="GI:1173624"
/translation="MAFVGLTRFIVSHKTAKEILSPAPADRPKESAVELLFNR
AMGAPPGDYWRNLRISSTYLSPREVSPKORSBIGMTRDMKMMERNGVEV
RMLHLYSLNNIMLVKKEKDEKLELILKEGYELGIFMWDHLLPLGLD
LOGVRRCTILVKNVFKKIIDHHRANGVIDEGEDFVDVLGLBEKDRSE
SDWAVLWEMIPECTDTVALILEMTLAPMLAPDIOSKOVETDVSRLPDLSDI
ORLPYLOSIVKERTRMHPPGLPPLSMARLAIHDVDPDGHNPACTTMMVMAITHEBC
NMAEPKFNDFRFTDEVDNLSDLRLAPGSKRVCPKGTMLAAVHMLAQLLKSF
KLPSRNGVLSBCLKMSLEKMLVCAVPRPE"
BASE COUNT 456 a 364 c 428 g 551 t
ORIGIN
Query Match 22.3%; Score 355; DB 8; Length 1799;
Best Local Similarity 58.0%; Pred. No. 2.8e-96;
Matches 744; Conservative 0; Mismatches 500; Indels 39; Gaps 5;
QY 304 TTCGAGCCCAAGAGCTCATGGCTTTCAGCTTAGAGAGAGCTGCGGTGATGTCAGGTGC 363
DB 273 TTGAGCCCTCCGCTCATGGCTTTCAGGTGCGCTCACTGCTTCATCGTTTAAAGC 332
QY 364 AATCCGACGTAGCGAAGAGATTCGATATAGCCCGGTTTTTGTATGACCGGTTAA 423
DB 333 CACCCGAAACCGGAAAGAGATTCCTTCAAGCCACCTTCGTATCGGCCCATTTAA 392
QY 424 GAATCGCTTACTCACTGATGTTTAAAGAGCAATGTTTGAACACACGCGTTTAC 483
DB 393 GAATCAGATACGAATCTCTGTTTAAATGCGCTATGCGTTTGGCCCATTTGGAGATTAC 452
QY 484 TGGCGAAGCTTGGCCGTATCGCTTGACACCATCTCTTATGACAAACAAATCAGAGA 543
DB 453 TGGAGAAACCTGAGAGAGATTCGTCCACATATCTTTCAGTCCGCGCAGATTCTCG 512
QY 544 GCCGAGACGCAACGACAGATGATCTCAAGCCAGATGGTGTGTTCTTGAAGAAACAGA-- 601
DB 513 TTCGAGAGCAACGAGATGATGGCGAAGAGATGTTGGGATATGAAAGAAATGATG 572
QY 602 -GTAGTAACGAACCTGTTTTTGTTCGTGAGTTCCTTAAACGCGCTGCTTAAACAATG 660
DB 573 GAGAGAAATGAGATTGTAAGATGAGAGATGTGCACTACGCGTCTTTGAATTAACATC 632
QY 661 ATGTGCTCTGATTCGACAGATGATGAGTGTAAAGAAACCACTGTGAGTTAGCTGAA 720
DB 633 ATGTGACTGTTTTGGGAAAGATTTGATTTTCCAAAGATGAGGGGTTGAGCTTGAG 682
QY 721 ATGTGC--GAAGAGATTATGATTTGCTCGGAAAGTGAATTGAGTATCACTTCT 777
DB 693 TTGATCTTAAAGAGATGATGATCTTGGATCTTCACTGGGATGATCACTTCT 752
QY 778 TGGCATGCGAGTTGATCTCTCAAGACTCCGCTGATGATTTCCACATCGTACCAAG 837
DB 753 CTTTGGAGATGATTTGCAAGGTGTGAGAGAGATGACAGAACCTTGTGGCTAAG 812
QY 838 GTAAACCGGTTTGTATCCCGATTATATCCGACACCGTATCAACCC----- 885
DB 813 GTCAATGATTTTGAAGAGATCATAGACAGCATAGAGAGCAACCGGCTAGGG 872
QY 886 -----GGTATTTGCCCTGATTTGCTGACGTTTTGCTCTCCCTCCATGATGAT 939

```

```

DB 873 ATTGATAGCGGTGAAGAGTGAAGATTTTGTGATGTCCTTCTGTTGGAGAGAAAGAT 932
QY 940 AATATATCCGACCCCGACATATATGCGCTTTTGGAGATGATATTCAGAGAAACAGAC 999
DB 933 AGACTCTCAGATCTGATATGATGTCAGATCTTTGGGAAAGATCTTTAGAGAACTGAT 992
QY 1000 ACAGTGGGCTTATATGAGTGAATCTCGCTGAGATGTCCTTCATCCAGATATGCAA 1059
DB 993 ACTGTGCGATCTATATGAGATGACGTTGGCTAGATGATCTTTCATCTCGATATTCAA 1052
QY 1060 TCACGCTACAAACAGCTGATCACTAGTATGAGGAAATCAAGACCTTAGATGATCT 1119
DB 1053 TCGAAGGACAAAGTATGATGATTCGTGCTGATGCTTCAAGGCCAGTATGGAATCT 1112
QY 1120 GACTTGGCTTCACTTCCATCTATCAAGCGCTGAGTGAAGAGATTTAGGCTTCATCT 1179
DB 1113 GATATCCAAAGCTTCTTATCTCCATCTATAGTAAAGAAACCTTCGATGATCTCT 1172
QY 1180 CCAGGCCCACTTCTATCATGAGGCCGCTTGGCCATTAACAGACAGATGATGATGCT 1239
DB 1173 CCGGGGCTCTATGTCATGAGGCTGCTAGCTATCCATGACGTTCTGTTGATGATGCTAC 1232
QY 1240 CTTGTTCCGGCAGGAGCAACAGCAATGATGAAATGTTGGGCCGTTATGCAATGATCCAC 1299
DB 1233 ATGATTTCTGCTGGGACGACCTGCAATGATGAAACATGTTGGCAATTAACATGACGAATGC 1292
QY 1300 GTGTGGGTGATCTTTGAGTTTAACTGAGAGTTCGTGGCAAAAGAGTGAAGTG 1359
DB 1293 AACTGGCTGAGCTTAACAATTCATCTGATGATTCATCGATGAAAGATG----- 1344
QY 1360 GAGTTTCGATCTTGGTGGATTTGACACTTGACCTTTCCGGTCCGATCGATTT 1419
DB 1345 ----TCAATATCTTGTTCGATTTAAGTGGGACCTTGGCTCCGGAAGAGATT 1400
QY 1420 TGGCCGGGAGAAATCTTGTATCTACCGTTATGTTTGGAGCGGATGATGATACAT 1479
DB 1401 TCCCTGGCAAAACATGATGATGCTGCAATTCATCTTGTGTTGCTCAATGCTGAAA 1460
QY 1480 GAGTTTGAATGGGACCGTCCGATGTAAGCGCTTGAATCTTGAAGAACTGAGCTT 1539
DB 1461 AGCTTCAATGCTTCTTCTC--GAGAATGATGATGATGCTGAGTGCCTTAAAGATG 1517
QY 1540 TCTTGCGAGATGCTAATCTCT 1562
DB 1518 TCTCTCGAGATGAGAAATCTTCT 1540

```

RESULT 14
 AF049067
 LOCUS 1990 bp mRNA linear PLN 06-MAR-1998
 DEFINITION Pinus radiata cytochrome P450 (PHE74) mRNA, complete cds.
 ACCESSION AF049067
 VERSION AF049067.1 GI:2935524
 KEYWORDS
 SOURCE
 ORGANISM
 Pinus radiata (Monterey pine)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
 1 (bases 1 to 1990)
 AUTHORSHIP Bishop-Hurley,S.L., Walter,C. and Gardner,R.C.
 TITLE Isolation and Expression of Abundant mRNAs during Somatic
 Embryogenesis of Pinus radiata
 JOURNAL Unpublished
 2 (bases 1 to 1990)
 REFERENCE Bishop-Hurley,S.L., Walter,C. and Gardner,R.C.
 DIRECT SUBMISSION Submitted (18-FEB-1998) Biotechnology Division, NZ Forest Research
 INSTITUTE, Sala Street, Rotorua, New Zealand
 LOCATION/Qualifiers
 1. 1990
 /organism="Pinus radiata"
 /mol_type="mRNA"
 /db_xref="taxon:3347"


```

/tissue_type="somatic embryonic"
/dev_stage="stage 1 embryo"
1..1990
/gene="PRB74"
3..1734
/gene="PRB74"
/codon_start=1
/product="cytochrome P450"
/protein_id="AAC05148.1"
/db_xref="GI:2935525"
/translation="MENRBSGGSGGMMVCLPFTKDPAYFLHSSDDVSNRMPV
VIALIVCAVIVSWLSPPGCAVAGHKGKRVLPKGMPIIGSLMDVSGV.PHRK
LESLARLHGAQKQMSFSGCTPAVITDPEVARELLSPHPANRPLKQSAQOLFGRA
IGFANNGITWRLRIRASAHLPAPRIIAHAGQAVVAMLDIQEYHSKVPR
RHLQALANIMGSVFGRRFDMSHENREKREKREVBGFOLDAEFMADHLPWLRPL
DPLRHARCARLVPRVTVFSNII EQRREQRRESDDQCFDVLVLSLQGEKDEE
DMIAVLMEIRFGIDTTLALTEMTWALVLEHBEAKQAKQELDVGHDSVKDSDIP
KLPIQAVKALRKHPPGLSWARLSTEDVNMGDGCVPAAGTAAVNMMSLTHDN
IMESPEERPRFVVEGGEVDVGRNDRSLAPGAGRVCPGKALGLATVNLVAVL
LHFEWLPFAHEPVDLSVTLKSCENARPLHCPVTVIVPAKESD"
BASE COUNT      408 a      472 c      560 g      550 t
ORIGIN
      21.9%; Score 348.2; DB 8; Length 1990;
      Query Match      55.5%; Pred. No. 3.3e-94;
      Best Local Similarity
      Matches 768; Conservative 0; Mismatches 588; Indels 27; Gaps 4;

211 ATCCCGGTCCAAAGGCTTCCCTTGGTGGAGACATGTCATCATGTCAGACACTTA 270
212 ATCCCGGTCCAAAGGCTTCCCTTGGTGGAGACATGTCATCATGTCAGACACTTA 270
307 ATTCCTGGCCCCAAAGATGGCCGATTAATGGAGTCTAATGAGATAGAGCGTGGCTT 366
271 GCTCACCGACGATGCGTGTGACGTGAGAAATTCGAGCGCAAGAGCGTCAATGCTTC 330
367 CCTCACCGACGATGCGTGTGACGTGAGAAATTCGAGCGCAAGAGCGTCAATGCTTC 426
331 AGCTTAGAGAGACTGCGGTGATCGTCACGTGCAATCCCGACGTGCGAAGAGATTTCG 390
427 AGCTTAGAGAGACTGCGGTGATCGTCACGTGCAATCCCGACGTGCGAAGAGATTTCG 486
391 AATAGCGCGGTTTTGGTGTGATGCAACGGGTTAAGATCGGCTTACTCATGATGTTAAC 450
487 AACTCCCTCACTTCCGCAACAGCCCTCAAGCAATCCGCAACAGCACTGCTGTTGGG 546
451 AAGACATTTGTTTGGACACACGCGTGTTAATCGCGAAGCGTTCGCCGATTCGCTTCG 510
547 AAGGCTATTGGCTTCGACCCCAACGGGCGCTACTGAGGCTGCTCAGAGAAATTCCTCT 606
511 AACCATCTTTTGTAGTCAAAACAAATCAAGAGCGGAGAGCGCAACGACGATGATCTCA 570
607 GCTCATCTCTTGGCTCTCGCAGATGCTGCCCATGAACTGGCCGCGCAGCGACCTT 666
571 AGCCAGATGTTGAGTTTCTTGA AAAACAGAGTGAAGAACCCCTGTTT---GTTCGT 627
667 GTGGCAATGCTGATGATATCCAAAGAGTATCATTTAAGGTTGTTTGGGTTAAGA 726
628 GAGTTCCTTAAACCGCGCTCGCTTACACATGATGCTGCTGATGTTGGAACAAGAT 687
727 CGACACCTGCAAGACCGCGCTGAAACAATATCATGAGGAGCGTCTTGGAGAGAGTTC 786
688 GAGCTTGA AAAAACCATGTTG-----AGTTAGTGAATGATCGAAGAGTTAT 738
787 GACATGCTCATGATGATGAAGAAAGTGAAGAGCTCGGAGATGTTGATGAAGGTTTC 846
739 GATTTCCTGGAACGTTGAATTTGATGATGATCACTTCTTGGCTATCGAAGTTTATCCT 798
847 CAACTGTATGAGACGTTCACTGCGGAGATCTCCATGCTGCGACCGCTGACCTT 906
799 CAAGAGCTCGGCTCATGATGTTCCACACTGTATCCAAAGGTTAAACCGGTTTATCCCG 858
907 CTTCGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
859 ATTATATCCGAACCGCTA-----ATCAACCGGTTATTTGCTGCTGATTTTC 906

```

```

Db 967 ATCATGCAAGACACCGCTCGGGAAGAGAGCGTGAAGAAAGCGGATCAATGATTTT 1026
Qy 907 GTGACGTTTTCCTCTCCCTCCATGTTCAATTAATATCCGACCCGACATATGCC 966
Db 1027 GTTGAAGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1086
Qy 967 GTTCTTTGGAGATGATATATTCAGAGAAACAGACAGTGGGCTTATGATGATGATC 1026
Db 1087 GTTCTTCTGGAATATATCTTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1146
Qy 1027 CTGCTGAGATGCTGCTTATCATGATATGCAATACGATACAGGTTCAAAAAGAGTGAATCAA 1086
Db 1147 ATGGCAGAACTGTTCTGCACTTCAAGCTTCAAGAAAGAGTCAAGAGAGTTGAAGCC 1206
Qy 1087 GTAGTGGGAGAAATCAAGAGCCCTAGATGATGATGATGATGATGATGATGATGATGAT 1146
Db 1207 GTAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1266
Qy 1147 GCTGTGTGAAAAGAGATATGAGGCTTATCTCTCAAGCCCACTTCTATCATGAGCCGCT 1206
Db 1267 GAGGTGTGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1326
Qy 1207 TTGGCCATTAACAG---CACATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1263
Db 1327 CTCTCTACAGAAAGATGATCAACATGAGGAGATGATGATGATGATGATGATGATGATGATGAT 1386
Qy 1264 ATGGTGAACATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1323
Db 1387 ATGGTGAACATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446
Qy 1324 AAACCTGAGAGGTTCTGTCGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1383
Db 1447 CGTCCGAGAGGTTTGTGTTGTTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1506
Qy 1384 TTGAGACTTGACCTTTCGGGTTGAGGCTGCTGAGATTTGCCCCGAGAGAGATTTGTTT 1443
Db 1507 CTGAGCTTCGCTGCTTTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1566
Qy 1444 ACTACGTTATGTTTGGACGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1503
Db 1567 GCCACTGTTATCTCGGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1626
Qy 1504 GGTAGCGGTTGACTTATCTGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1563
Db 1627 GAACATCGGTTATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1686
Qy 1564 CCT 1566
Db 1687 CAT 1689

RESULT 15
AL772960/c 565 bp DNA 1linear STS 11-JUN-2003
LOCUS Arabidopsis thaliana transposon insertion STS SM_3_556, sequence
DEFINITION
AL772960
ACCESSION AL772960
VERSION AL772960.1
KEYWORDS STS; STS; sequence tagged site.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B.,
Murphy, G., Jangnam, S., Leary, C., Jones, J.D.G. and Bevan, M.
JOURNAL Unpublished
AUTHORS Clarke, J.H.
TITLE Direct SubMISSION
JOURNAL Submitted (18-JUN-2002) Clarke J.H., John Innes Centre, Colney
Lane, Norwich, NR4 7UJ, UK

```

COMMENT

AT denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, GT a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SW a defective suppressor mutator transposon. 3 denotes a sequence derived from the 3' end of the transposon, 5 denotes a sequence derived from the 5' end of the transposon. BBSRC GARNET, ATIS project
On-line seed stock requests: <http://naec.mot.ac.uk/NAASCstockcode: N56511>.

FEATURES

source

Location/Qualifiers
1..565
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia-0 NASC stock code N1092"
/db_xref="taxon:3702"
/clone="AL138642"
/note="Derived from superpool 12.34 NASC code Unknown"

STS

BASE COUNT 167 a 151 c 112 g 135 t
ORIGIN

Query Match 21.1%; Score 336.2; DB 11; Length 565;
Best Local Similarity 83.9%; Pred. No. 1.3e-90;
Matches 380; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

```
QY 975 GGAGATGATATTCAAGAGAACAGACAGATGCGGCTTAAATCGAGTGAATCTCGCTAG 1034
    |||
Db 453 GGAAATGATATTCAAGAGAACAGACAGATGCGGCTTAAATCGAGTGAATCTCGCTAG 394
    |||

QY 1035 GATGTCCTTCATCCAGATATGCAATCAACGCTGATCAAGATCAAGATCGG 1094
    |||
Db 393 GATGTCCTTCATCCAGATATGCAATCAACGCTGATCAAGATCAAGATCGG 334
    |||

QY 1095 GAATCAAGAGCCCTAGATGATCTGACTTGCTTCACTTCATATCAACGCTGTGT 1154
    |||
Db 333 ACGATCAAGGCTGTGAGAGATGATGAGTGTCTAGATATCTAACGCTGTGT 274
    |||

QY 1155 GAAAGAGATGAGGCTTCATCTCCAGGCCCACTTATCATAGGCGCGTTGGCCAT 1214
    |||
Db 273 GAAAGAGATGAGGCTTCATCTCCAGGCCCACTTATCATAGGCGCGTTGGCCAT 214
    |||

QY 1215 AACAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1274
    |||
Db 213 CAAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 154
    |||

QY 1275 GTGGGCGGTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1334
    |||
Db 153 GTGGGCGGTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 94
    |||

QY 1335 GTTCGTGCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGC 1394
    |||
Db 93 TTTGTAGCCAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGC 34
    |||

QY 1395 ACCTTTCGGGTCGGGTCGGGTCGGGTCGGGTCGGGTCGGGTCGGGTCGGGTCGG 1427
    |||
Db 33 ACCGTTCCGGTCGGGTCGGGTCGGGTCGGGTCGGGTCGGGTCGGGTCGGGTCGG 1
    |||
```

Search completed: January 17, 2004, 03:16:35
Job time : 5890 secs

Tue Jan 20 17:43:17 2004

us-10-022-025a-1.rn1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: January 16, 2004, 09:48:47 (Search time 124 Seconds
(without alignments)
7151.119 Million cell updates/sec

Title: US-10-022-025a-1

Perfect score: 2009

Sequence: 1 aataaataaataaatacgtta.....atagagattagagatttt 2009

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 563978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 113956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	420.8	20.9	1611	3	Sequence 7, Appl
2	200.4	10.0	1799	2	Sequence 3, Appl
3	123.2	6.1	1704	3	Sequence 17, Appl
4	101.2	5.0	1708	3	Sequence 1, Appl
5	98.2	4.9	275	4	Sequence 2391, Ap
6	89	4.4	1788	3	Sequence 9, Appl
7	88.8	4.4	1893	3	Sequence 1, Appl
8	84.6	4.2	2174	3	Sequence 63, Appl
9	84.6	4.2	2174	3	Sequence 63, Appl
10	83.2	4.1	1761	3	Sequence 3, Appl
11	83.2	4.1	1927	3	Sequence 64, Appl
12	83.2	4.1	1927	3	Sequence 64, Appl
13	81.4	4.1	5156	2	Sequence 3, Appl
14	81.4	4.1	5156	2	Sequence 3, Appl
15	79.2	3.9	1722	3	Sequence 1, Appl
16	79.2	3.8	1755	1	Sequence 29, Appl
17	77.2	3.8	1755	1	Sequence 29, Appl
18	77.2	3.8	1755	1	Sequence 29, Appl
19	76.4	3.8	1812	1	Sequence 37, Appl
20	76.4	3.8	1824	3	Sequence 1, Appl
21	76.4	3.8	1824	3	Sequence 1, Appl
22	75.6	3.8	1835	4	Sequence 232, App
23	75.6	3.7	1835	4	Sequence 5, Appl
24	75	3.7	1838	2	Sequence 1, Appl
25	75	3.7	1838	2	Sequence 1, Appl
26	73	3.6	1620	4	Sequence 11, Appl
27	73	3.6	1620	4	Sequence 11, Appl

28	72.8	3.6	543	4	US-09-615-192A-230	Sequence 230, App
29	71.8	3.6	1894	4	US-09-564-808-3	Sequence 3, Appl
30	70.2	3.5	1657	3	US-08-948-564-11	Sequence 11, Appl
31	69.2	3.4	1806	4	US-09-351-229-3	Sequence 3, Appl
32	68.6	3.4	1880	4	US-09-564-808-1	Sequence 1, Appl
33	66.6	3.3	1781	4	US-09-499-302A-1	Sequence 1, Appl
34	65.4	3.3	1812	1	US-07-912-300-28	Sequence 28, Appl
35	65.4	3.3	1812	1	US-08-285-309-28	Sequence 28, Appl
36	65.4	3.3	1812	2	US-08-502-046-28	Sequence 28, Appl
37	64	3.2	1506	3	US-09-158-767-7	Sequence 7, Appl
38	64	3.2	1506	3	US-09-158-767-8	Sequence 8, Appl
39	64	3.2	1506	3	US-09-158-767-9	Sequence 9, Appl
40	64	3.2	2261	3	US-09-158-767-1	Sequence 1, Appl
41	63	3.1	1737	4	US-09-126-420A-1	Sequence 1, Appl
42	63	3.1	1824	3	US-08-948-564-13	Sequence 13, Appl
43	62	3.1	386	1	US-08-313-075A-55	Sequence 55, Appl
44	61.4	3.1	1634	4	US-09-126-420A-2	Sequence 2, Appl
45	61.4	3.1	1644	3	US-08-948-564-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-948-564-7
Sequence 7, Application US/08948564
Patent No. 6121512
GENERAL INFORMATION:
APPLICANT: Simitsky, Balazs
APPLICANT: Dewey, Ralph E.
TITLE OF INVENTION: Cordin, Frederick T.
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESS: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512th Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,564
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCE/DOCKET NUMBER: 5051-409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-854-1400
TELEFAX: 919-854-1401
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1611 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 20..1588
US-08-948-564-7
Query Match 20.9%; Score 420.8; DB 3; Length 1611;
Best Local Similarity 58.1%; Pred. No. 4.9e-107;
Matches 1029; Conservative 0; Mismatches 487; Indels 256; Gaps 6;

Tue Jan 20 17:43:17 2004

us-10-022-025a-1.rni

Page 2

QY 221 AGCCAAACCAACCTTGGCTTCTCCCTCCCTGCGCGTCAATCATCTGGCTGGCATATCT 280
Db 74 ACCCAAAAACCTTGATGAGGTCTTTTATGATGAGGCTCACTCTGTAAACATGACT 133
QY 281 CTCTCTTATGAGCACTATCCCGGTGACCTGCTGGGGGAAATACCTCTTCCGCGGTTA 340
Db 134 TTTATTAATGCTACACCCCGGTGCTGCTGGGGGAAATACCTCTTCCGCGGTTA 186
QY 341 ATATCGGTTCAATCAAAACCGGAAACGTTATTCGCGGTCCAAAGGCTTCCCTTGTG 400
Db 187 CTCTCCCGCTTCAATCAAAACCGGAAACGTTATTCGCGGTCCAAAGGCTTCCCTTGTG 235
QY 401 GGAAGCATGCTCATGCTGTAAGCACTGTAAGTCAACGAAATGCTGATGAGCTGAG 460
Db 236 GGAAGCATGAGGCTCATGACTTCC--CTGACCATATACGTTATGCAAGCGCGCGGCC 282
QY 461 AAATTCGAGCCAAAGGCTCATGCTTCAAGCTTGAAGAGACTGCGGTGATGCTCAAG 520
Db 293 ACATGCAAGCCAAAGGCTCATGCTTCAAGCTTGAAGAGACTGCGGTGATGCTCAAG 352
QY 521 TGCAATCCGACGAGTGAAGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 580
Db 353 TGCAATCCGACGAGTGAAGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 412
QY 581 AAAGATCGGCTTCAATCAAAACCGGAAACGTTATTCGCGGTCCAAAGGCTTCCCTTGTG 640
Db 413 AAAGATCGGCTTCAATCAAAACCGGAAACGTTATTCGCGGTCCAAAGGCTTCCCTTGTG 472
QY 641 TACTGGGAGAGCTTCCGCGTATGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 700
Db 473 TACTGGGAGAGCTTCCGCGTATGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 532
QY 701 AGAGCCGAG 760
Db 533 GCTCTGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 589
QY 761 AGATGTAAGCAACCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 820
Db 590 AAGCGCCAG 649
QY 821 ATGCTCTGATATTCGAGCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 871
Db 650 ATGCTCTGATATTCGAGCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 709
QY 872 TTAAGTGAATGCTGAG 931
Db 710 CTGGAATATTAAGTGAAG 769
QY 932 CTTCCTTGGCTATGAG 991
Db 770 CTTCCTTGGCTATGAG 829
QY 992 CCAAGAGTAAACCGGTTGTATCCGAGTATATCCGAGTATATCCGAGTATATCCGAGTAT 1051
Db 830 CCAAGAGTAAACCGGTTGTATCCGAGTATATCCGAGTATATCCGAGTATATCCGAGTAT 889
QY 1052 TTGCTCTGATATTCGAGCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1111
Db 890 ACCAATCGATTTTGTGAG 949
QY 1112 CCGACATATGCGCGTCTTGTGGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1171
Db 950 TCCGACATGATGCTGACTTTG-----TCAATCCGAGTCCCAAGGCTTCCCTTAT 972
QY 1172 TATATTTTGTAAAAATGCTTGAATATTAAGATTAATAAATGAGTAAATTTGAGCTCAAT 1231
Db 973 ----- 972
QY 1232 GTGAGGTGTTGAGATATTAAGAGATTAAGAGATTTTGTGATGATCAAAATTAGTAA 1291
Db 973 ----- 972
QY 1292 ATTTAGATTTTAAACAGTTTCTATATAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1351

Db 973 ----- 972
QY 1352 ATTAAGAGATGATATTAAG 1411
Db 973 -----GGAATGATATTAAG 1026
QY 1412 CCGTATGATGCTCTTATTCAGATATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1471
Db 1027 CCGAGAGTGGGCTTCACTTCACTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1086
QY 1472 AGTGGGAAATCAAG 1531
Db 1087 TGTGGAAAG 1146
QY 1532 TGTGGTAAAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1591
Db 1147 GGTGGTAAAG 1206
QY 1592 GGGCATAACAG 1651
Db 1207 GTCCATCATGATGAG 1266
QY 1652 GAAATGAG 1711
Db 1267 CAACAGTGGGCTATTTGAG 1326
QY 1712 TGAGAGTTCGAG 1771
Db 1327 CAGAGAGTTCGAG 1386
QY 1772 ACTTGCACCTTTCGAG 1831
Db 1387 ACTTGCACCTTTCGAG 1446
QY 1832 CGTTATGTTTGAAG 1891
Db 1447 GGTGAATTTTGAAG 1506
QY 1892 CCGCGTGAATTAATGAG 1951
Db 1507 GGTGTTGATGAG 1566
QY 1952 TAAATTTGCGCGTGAAG 1983
Db 1567 CAAAGTGGCGCGAG 1598

RESULT 2
US-08-560-398-3
Sequence 3, Application US/08560398
Patent No. 5907082
GENERAL INFORMATION:
APPLICANT: O'Neil, Sharmen
APPLICANT: Nadeau, Jeanette
TITLE OF INVENTION: Ovine-Specific Gene Expression
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,398
FILING DATE: 17-NOV-1995
CLASSIFICATION: 435

Page 4

```

RESULT 4
US-08-991-677-1
Sequence 1, Application US/08991677A
Patent No. 6252135
GENERAL INFORMATION:
APPLICANT: Chiang, Vincent L
APPLICANT: Carraway, Daniel T
APPLICANT: Smeltzer, Richard H
TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
FILE REFERENCE: 50617
CURRENT APPLICATION NUMBER: US/08/991,677A
CURRENT FILING DATE: 1997-12-16
EARLIER APPLICATION NUMBER: US 60/033,381
EARLIER FILING DATE: 1996-12-16
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1708
TYPE: DNA
ORGANISM: Liquidambar styraciflua
FEATURE:
NAME/KEY: CDS
LOCATION: (48)..(1571)
US-08-991-677-1

```

Db 1334 GCCCTTTGGTGCAGGAGGCGTCTTTTCCCGGTCACAACTTGGCATCAATTGGTTCAC 1393
 Oy 1838 GTTTTGCACGCGCATGATGTTACATGAGTTGAATGGGACCGTCCGATGCT 1889
 Db 1394 ATCATATAGGGTACCTATTGGACCACTTTCTATGGAGCCCTCTAAAGGT 1445

Query Match	4.9%	Score 98.2;	DB 4;	Length 275;
Best Local Similarity	64.7%	Pred. No. 9,1e-18;		
Matches 178; Conservative	0;	Mismatches 93;	Indels 4;	Gaps 2,

RESULT 6
US-08-948-564-9

Patent No. 6121512
GENERAL INFORMATION:
APPLICANT: Siminszky, Balazs
APPLICANT: Dewey, Ralph E.
APPLICANT: Corbin, Frederick T.
TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
METHODS OF SEQUENCES: 23
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512ch Carolina
COUNTRY: USA

Tue Jan 20 17:43:17 2004

us-10-022-025a-1.rtf

Page 5

```

? ZIP: 27627
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent in Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/948,564
? FILING DATE:
? CLASSIFICATION: 800
? ATTORNEY/AGENT INFORMATION:
? NAME: Bennett, Virginia C.
? REGISTRATION NUMBER: 37,092
? REFERENCE/DOCKET NUMBER: 5051-409
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 919-854-1400
? TELEFAX: 919-854-1401
? INFORMATION FOR SEQ. ID NO.: 9:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1788 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 6..1601
? US-08-948-564-9

```

Query Match	4.4%	Score 89	DB 3	Length 1788
Best Local Similarity	48.2%	Pred. NO. 8e-15		
Matches 285	Conservative	0	Mismatches 300	Indels 6
			Gaps	1,1
QY	1356	CAGAGATGATTTAGAGGAAACAGACACAGTTGGCTTCTTAATCGATGAGATCTCGCT	1415	
DB	972	CTGAATCTGATTTTAGCAGGAAGCAGCACCCACATGATTTCACTAACATGCGGTATCT	1031	
QY	1416	AGGATGGTCTTCATCCAGATATGGAATCAACGGATTAACAAACGAGCGGATCAAGATGC	1475	
DB	1032	CTGCTACTTAACCATTCATATGGAATCAAAAAAATGCCAATGATATGACACTATATTT	1091	
QY	1476	GGGAATCAAGACCCCTGATGATCTGACTTGGCTTCACTTCATATCTAAACGCTGTG	1538	
DB	1092	GGAGAGACAGAAAGGTGAAAGAAATCAGCACTAACAAAGTTGGTATACCTCCAGGCATT	1151	
QY	1536	GTTAAAGATATTAGAGCTTCATCTCCAGGCCCATCTTCATCATGAGGCCGTTTGGCC	1595	
DB	1152	GTTGAAGAAACAATGGGTGTATTCACCAAGTCTCTTATCAACCTGTGTGAGCGCATG	1211	
QY	1596	ATTAACAGACCATGCTTGTATGTGTCTTTTGTCCGACAGGACACAGCATGTGTAAAC	1655	
DB	1212	GAAGACTGACCTTCTCAGGTGGCTATACATTTCTCTGAGACAGTTTAATGGTAAT	1271	
QY	1656	ATGTGGCCGATATCCATGATATCCACACGTGGGGTTATGCTTTGAGATTAAACCTGAG	1715	
DB	1272	GCTTGAACATCCACCGGAGTGTGTGTGTTTGAAGTATCTCATGATTTCAAGCCTGGA	1331	
QY	1716	AGGTTGTGGCAAAAGAAAGGTGAGATTTTCGGTTCTTGAGTCCGATTTTGAACATT	1777	
DB	1332	AGGTTTTCACAAAGCCACAAAGATGTTGATGTGAAGGATC-----AGAACTATGAGCTC	1386	
QY	1776	GCACTTTGGGTCGGTCTGTCGATTTTCCCCCGGAAAGATCTTGTGTTTACTACCGTT	1833	
DB	1386	GTCCTTTTGGTTCTGGAAGAGAGACATGCCCTGGAACCTCGCTGCTCGCTGTHGATG	1445	
QY	1836	ATGTTTGGACGCGCATGATGTTTAATGATGTTTAATGAGGAGCCGTCCGATGAGTAAACGCG	1899	
DB	1446	CACCTTACCATGGCTGACTGTTTAATCTTTTCAATGTGCTTCTCTTCAATCAAGTT	1505	
QY	1896	GTTGACTTATCTGAGAAACTGAGGCTTTCTTGCAAGATGGCTATCTCTTT	1946	
DB	1506	GTCGACATGACAGAGGACATTTGAGATCTCCAAATTTTAAAGAACCCCGCTT	1556	

```

1      RESULT 7
2      US-08-532-065B-1
3      Sequence 1, Application US/08532065B
4      Patent No. 5733507
5      GENERAL INFORMATION:
6      APPLICANT: Ohta, Daisaku
7      APPLICANT: Mizutani, Masaharu
8      TITLE OF INVENTION: Plant Geraniol/Nerol 10-Hydroxylase and
9      TITLE OF INVENTION: DNA Coding Therefor
10     NUMBER OF SEQUENCES: 6
11     CORRESPONDENCE ADDRESS:
12     ADDRESSEE: No. 5753507artle Corporation
13     STREET: 59 Route 10
14     CITY: East Hanover
15     STATE: NJ
16     COUNTRY: USA
17     ZIP: 07936
18     COMPUTER READABLE FORM:
19     MEDIUM TYPE: Floppy disk
20     COMPUTER: IBM PC compatible
21     OPERATING SYSTEM: PC-DOS/MS-DOS
22     SOFTWARE: PatentIn Release #1.0, Version #1.30
23     CURRENT APPLICATION DATA:
24     APPLICATION NUMBER: US/08/532.065B
25     FILING DATE: 22-SEP-1995
26     CLASSIFICATION: 800
27     ATTORNEY/AGENT INFORMATION:
28     NAME: Meigs, J. Timothy
29     REGISTRATION NUMBER: 38,241
30     TELECOMMUNICATION INFORMATION:
31     TELEPHONE: 919-541-8587
32     TELEFAX: 919-541-8688
33     INFORMATION FOR SEQ. ID NO.: 1:
34     SEQUENCE CHARACTERISTICS:
35     LENGTH: 189 base pairs
36     TYPE: nucleic acid
37     STRANDEDNESS: single
38     TOPOLOGY: linear
39     MOLECULE TYPE: cDNA
40     ORIGINAL SOURCE:
41     ORGANISM: Arabidopsis thaliana
42     IMMEDIATE SOURCE:
43     CLONE: P450-4
44     FEATURE:
45     NAME/KEY: CDS
46     LOCATION: 6..1490
47     US-08-532-065B-1

```

	Query Match	4.4%	Score 88.8	DB 1	Length 1893
	Best Local Similarity	50.5%	Pred. No. 9.4e-15		
	Matches 278	Conservative	0	Mismatches 257	Indels 15
				Gaps	2
Qy	1340	TTTTAGATTATATTACAGAGATGATATTCAGAGAAACAACAAGTTGCGGCTTTAT	13399		
Db	905	TATGAAACACCTTCCTTGATATGTTTACAGACGACAGATACAGAGCTTAGACCT	964		
Qy	1400	CGAGTGAATCCCGGTAGATGTCCTTCATCAGATATGCAATCAACGGTACAAACGA	1459		
Db	965	GGAGTGGCCATGACAGAGTTACTTTAAAACTTAAAAAGATGGCAAAAGCTCAGCGCGA	1024		
Qy	1460	GCTGATCAAGTAGTCGGGAATCAAGAGCCCTAGATGAATCTGACTTGGCTTCACTTC	1519		
Db	1025	GATCGATTGTGTAAGAGTCAAAACGGTATCGTTGAAGAGTCAGATATCTCAACCTCC	1084		
Qy	1520	AATATCAACGGGTGGTGAAGAAGATTGAGGCTTCACTCCACAGGCCACTCTATC	1579		
Db	1085	GATTTTCAAGCAGGTCGTGAAGAAATTTCCGGTTACATACGCCGTCTCCGCTTTAT	1144		
Qy	1580	ATGGGCGCTTTGGCCATTAACAGACAGATCGTTGATGTCGCTCTTGTCCGGCAGGAC	1639		
Db	1145	CCCG---CGAAAGGCCAGATCCGATGGGAGATCTCTGTTTCAATGAGCTCTTAAAGATC	1201		

SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/616,990
FILING DATE: 14-JUL-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Perity, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 2174 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Eustoma russellianum
FEATURE:
NAME/KEY: CDS
LOCATION: 92 to 1621
IDENTIFICATION METHOD: by experiment
SEQUENCE DESCRIPTION: SEQ ID NO: 63
US-09-616-990-63

Query Match 4.2%; Score 84.6; DB 3; Length 2174;
Best Local Similarity 49.4%; Pred. No. 1.5e-13;
Matches 306; Conservative 0; Mismatches 304; Indels 9; Gaps 3;

QY 1338 ATTTTGGATTATATTACAGAGATGATATTCAGAGAAACAGACACAGTGGCGTCTTA 1397
DB 971 AATATCAAGGCTCTTTATGAAACATGTATACCGGTACGGATACATCTTCACACGCTC 1030
QY 1398 ATCGAGTGGATCCCTCGTAGATGTCCTTCATCCAGATATGCAATCAACGGTACAAAC 1457
DB 1031 ATAGAGTGGCGCTGCGCGAGGTGCTTAAAGATTCATCTTAAGACAGGCCCAAGAA 1090
QY 1458 GAGCTGATCAAGTAGTGGGAAATCAAGACCCCTAGATGATCTGACTTGCGTCACTT 1517
DB 1091 GAATGACGCGTGTGATCGGCGAGACCGCGGTTCTTAGGCAAGCATATCAAAAGTTG 1150
QY 1518 CCATATCTAACCGCTGTGGTGAAGAAAGATATGAGGCTTATCTCCAGGCCCATCTTA 1577
DB 1151 CCATATCTCAAGCCATCTGCAAGAAAGCTTTGCAAGAGATCTTTC--CAGCGCTTTA 1207
QY 1578 TCATGGGCGCGTTGGCCATTAACAGACACAGATCGTTGATGCTGTTGTTCCGAGAGG 1637
DB 1208 AATCTCCACGAATCGGTGCGCAAGCATGGAAGTAAATGAGACATACCAAAAGGCG 1267
QY 1638 ACCACAGCATGTGAACATGTGGCCGATCGCATATCAACAGTGTGGGTGATCTT 1697
DB 1268 ACTGGCTCAGCGTTAAACATATGGCTATGGAAGAGATCCATCTGTGTGGAAATCCA 1327
QY 1698 TTGAGTTTAAACCTGAGAGGTTGCTGCAAAAGAGTGAAGGAGGTTTGGTCTT 1757
DB 1328 AATGAATTTAACCTGTATAGTCTTTTGAAGCAAGAAATGCA---AGATGATTCACGA 1384
QY 1758 GGGTCGATTTGAACATTTGACCTTTGGGCTGCGGTGTGAGATTTGCCCGGAGAAAT 1817
DB 1385 GGAATATGATTTGAGCTGATCCCATTTGAGCTGGAAGAAATTTCCGTGAAACAAG 1444
QY 1818 CTGTGTTTATACCGTATGTTTGAACGCGCATGATGTAACATGATGTTTGAATGGGA 1877
DB 1445 TTGGGAATATCTTATAGTGAATATATTTTGGAACTTTGTGATCTTTTGTGGGA 1504
QY 1878 ---CCGTCCGATGGTAAACGCGGTGACTTATCTGAAGAACTGAGGCTTTCTTGGAGATG 1934
DB 1505 TTGCATCTCTCTGATGATGAACATTAACATGATGATGATCTTTGGGCTTGTCTGCAAG 1564

QY 1935 GCTATCTCTCTGCTA 1953
DB 1565 GCAGTGCCTCTGCTGCTA 1583

RESULT 10
US-09-033-055A-3
Sequence 3, Application US/09033055A
Patent No. 6069241
GENERAL INFORMATION:
APPLICANT: OKAWA, HIDEO
APPLICANT: IMATSHI, HIROMASA
TITLE OF INVENTION: NOVEL CYTOCHROME P450 GENE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,055A
FILING DATE: MARCH 2, 1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PAUL E. WHITE, JR.
REGISTRATION NUMBER: 32,011
REFERENCE/DOCKET NUMBER: 9437/251563
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3651
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1761
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-033-055A-3

Query Match 4.1%; Score 83.2; DB 3; Length 1761;
Best Local Similarity 49.0%; Pred. No. 3.3e-13;
Matches 282; Conservative 0; Mismatches 280; Indels 6; Gaps 2;

QY 1351 TATTACAGAGATGATATTCAGAGAACAGACACAGTGGCGTCTTAATGAGTGAATCC 1410
DB 962 TATTACAGATGATGATCTGACAGCAAGATACATCTGCTGTCAACCAAGATGGCCAA 1021
QY 1411 TGGCTAGATGCTCTTATCCAGATATGATCAATCAACGTTACCAAGCTGATCAAG 1470
DB 1022 TGGCTAGATGATTAAGATCAATCAATGCTTCAAGAGATCCAAAGAACTGATATG 1081
QY 1471 TAGTGGGAATCAAGACCCCTAGATGATCTGACTTGGCTTCACTTCAATATCTAAG 1530
DB 1082 TTGTTGATTCGACCGTATGATACCGATCCGATCTGTTGATCTCAAGTACCTTGGT 1141
QY 1531 CTGTGTGAAGAAATGATTTAGGCTTCACTCCAGGCCCACTTATCATGAGGCCGT 1590
DB 1142 GTGTAGTACGTAACCATTTGAAATGCAACCTGCTGTCTTTCT---AATTCACATG 1198
QY 1591 TGGCATTAACAGACAGATCGTTGATGCTGCTTTTCCGACAGGACCAACGAATGG 1650
DB 1199 AATCAATTCGATATCAAGATCAACGGGTATATATCCAGCAAGACACGTGCTTCA 1258
QY 1651 TGAACATGTGGCCCGTATCGATGATTCACACAGTGTGGTATCTTGAATTAAC 1710
DB 1259 TCAACACATGCTCTGTGTAAGAAACAAAGATATGGAATATGATGATGATTTAGGC 1318

QY 1711 CTGAGAG---CTTCGTGGCAAAAGAGGTGAGTGGAGTTTGGGTTCTTGGGTCGAT 1767
DB 1319 CAGAGAGACATTGGCTTGCACATGATTAAGTAAATGATGATGAGGCGATTT 1378
QY 1768 TGAAGCTTGACCTTTGCGGTCGCGTCGATTTGCCCCGGAGAGATCTTGGTTTA 1827
DB 1379 TCAAGATTTTACCATTTAGTGTCTGAGAAAAGAGAGTGTCTGCGACCATTTAGAGTGA 1438
QY 1828 CTACCGTTATGTTTGGACGCGGATGATGTTACATGATTTGATGGGACCGTCGATG 1887
DB 1439 AATTGGTACTTATGCTTGGCTGATGATTTGCTTGTCTTGGAGTGGAGCCACAGATG 1498
QY 1888 GTAAAGCGCTTGACTTATCTGAGAACTGAGGCTTT 1923
DB 1499 GATTACGGCTGAGATATGTACACATTTAGGATTT 1534

RESULT 11

US-08-606-505B-64
Sequence 64, Application US/08606505B
Patent No. 6114601

GENERAL INFORMATION:

APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: KIYOKAWA, Shigeto
APPLICANT: SHIMADA, Yukihisa
APPLICANT: OHBAYASHI, Masaya
APPLICANT: SHIMADA, Ritsuko
APPLICANT: OKINAKA, Yasushi
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.

COMPUTER: IBM PS/VS
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/606,505B
FILING DATE: 23-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Perry, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200
SEQUENCE CHARACTERISTICS:
LENGTH: 1927 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE: Campanula medium
FEATURE:
NAME/KEY: CDS
LOCATION: 180 to 1748
IDENTIFICATION METHOD: by experiment
US-08-606-505B-64

Query Match 4.1%; Score 83.2; DB 3; Length 1927;
Best Local Similarity 49.5%; Pred. No. 3,4e-13;
Matches 273; Conservative 0; Mismatches 273; Indels 6; Gaps 2;

QY 1337 AATTTTGGATTAATTATTAACAGAGATGATATTCAGAGAAACAGACAGTTGGGTCCT 1396
DB 1097 AATGTTAAGGACACTCTTTTGGATTTATTCACGGCGGGGACAGATCATCATCAAGTGT 1156
QY 1397 AATCGATGATCTCTCGTAGGATGTCCTTCATTCAGATTTGCATATCAAGGTACAAA 1456
DB 1157 GATCGAATGGGACCTAGCGCAAAATGTTAAGCATTCAGACATCTGATACAGAACTT 1216
QY 1457 CGAGCTGATCAAGTATCGGGAATCAAGACCCCTAGATGATCTGACTTGGTTCACT 1516
DB 1217 AAAAAAGACAGATCATTTGGCAGAAACAGAAAGATTAAGCATCTGATACAGAACTT 1276
QY 1517 TCCATATCTAACGCTGTGTGTAAGAAAGATTTGAGGCTTCCTCCAGGCCCATCTT 1576
DB 1277 GCCATATTTCCAAAGCCATATGCAAGAAACATTCGAAACACCCCTT---CAGGCCCTT 1333
QY 1577 ATCATGGGCGCGTTTGGCCATTAACAGACAGATCGTTGATGTCGTCCTTTCGCGCAG 1636
DB 1334 AATCTCCCAAGATCTCAACAGAGCATGTGAAGTGAACGATTTCAATACCAAAAAA 1393
QY 1637 GACCAAGCAATGTGAACATGTGGCCGATATCGCATATCCACAGTGTGGTTGATCC 1696
DB 1394 CACTAGACTAATAGAGACATATGGGCAATAGGGAGGAGACCTTAAGTGTGGAAAAATCC 1453
QY 1697 TTTGAGTTTAACTGAGAGGTTGTCGCAAAAAGAGGTAGGTTGGATTTGGTTCT 1756
DB 1454 ATTAGATTTTACTCGGAACGTTTCTTGATGTAAGAAC---ACGGAATAATGATCCGC 1510
QY 1757 TGGGTCGATTTGAGACTTGACCTTTCGGGTGCGGTGTCGATTTGCCCGGAGAA 1816
DB 1511 AGGTATATCAATTTGATTTATCCATTTGGGGCTGAGCAAGATATGTGACAGGGCTAG 1570
QY 1817 TCTGTTTACTACCGTATGTTTGGACGCGCATGATTTACATGATTTGATGGGG 1876
DB 1571 AATGGAGGGCGCTGCGTGGATGATATTAGTACATTTGTCACCTATTTGATTGAA 1630
QY 1877 ACCGTCGATGG 1888
DB 1631 ATTGCTGATGG 1642

RESULT 12.

US-09-616-990-64
Sequence 64, Application US/09616990
Patent No. 6232109

GENERAL INFORMATION:

APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: KIYOKAWA, Shigeto
APPLICANT: SHIMADA, Yukihisa
APPLICANT: OHBAYASHI, Masaya
APPLICANT: SHIMADA, Ritsuko
APPLICANT: OKINAKA, Yasushi
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/VS
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/616,990
FILING DATE: 14-Jul-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992

Tue Jan 20 17:43:17 2004

us-10-022-025a-1.rn1

Page 9

ATTORNEY/AGENT INFORMATION:
NAME: Peirry, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 64 :
SEQUENCE CHARACTERISTICS:
LENGTH: 1927 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Campanula medium
FEATURE:
NAME/KEY: CDS
LOCATION: 180 to 1748
IDENTIFICATION METHOD: by experiment
SEQUENCE DESCRIPTION: SEQ ID NO: 64

Query Match	Similarity	4.1%	Score 83.2	DB 3	Length 1927
Best Local	49.5%		Pred. NO.3.4e-13		
Matches 273	Conservative	0	Mismatches 273	Indels	Gaps 2
QY	1337	AATTTTGGATTATTTATTACAGAGATGATATTCAGAGAAACAGACAGTTGGCGTCT	13366		
Db	1097	AAATGTTAAGCACACTCTCTTTGGATTATTTACAGCGCGGCGACGGATACATCAAGTGT	11566		
QY	1397	AATCGAGAGGATCCCGGTAGATGTCTTCATCCAGATATGCAATCAACGATACAAA	14566		
Db	1157	GATCGAATGGGCACTACGCCGAATGTTGAACCATGACAGATCTCAACCGGGCCACGA	1216		
QY	1457	CGAGCTGATCAAGTAGTCGGGAAATCAAGACCCCTAATGATATTTGACTTGGCTTCACT	1516		
Db	1217	AGAAATGACCAAGCATTTGGCAGAAACAGAAAGATTAACACATTTGACATACCAACTT	1276		
QY	1517	TCCATATTTAACGGCTGTGTGAAGAAAGTATTTGAGCTTCATCTCCAGGCCCACTTCT	1576		
Db	1277	GCATATTTCCAGGCATATGGAAGAAAGAAACATTCGAAAAACACCTTC---CAGCCCTT	13333		
QY	1577	ATCATGGGCGCGTTTGGCCATPACAGACACGATCTGTATGTGTCTTTGTTCCGACAG	1636		
Db	1334	AAACCTCCCAAGAACTCTCAACGAAAGCATGTGAATGACGGATTTCCATPACAAAAA	13939		
QY	1637	GACCAACAGAAATGTGAACATGTGGCGCGTATCGCATATCCACACGTGTGGTGTATCC	1689		
Db	1394	CACATGACTAATATGTAACATATGTGGCAATAGGAGGAGACCTTAAGTGTGGAAAAATCC	14531		
QY	1637	TTTGGAGTTTAACTGTAGAGGTTTGTTGGCAAAAAGATAGAGTGAAGTTTCCGTTCT	1756		
Db	1454	ATTGATATTTACTCCGAAACGTTTCTTGAATGAAAAAC---ACGGAAAAATTGATCCGG	1510		
QY	1757	TGGGTCGATTTTGAGACTTGACCTTTCCGGTCCGCTCGTCGGATTTGCCCGGAGAGA	1816		
Db	1511	AGGTATATCTTTTGAATTAATCCCATTTGGGGCTGAGAGAAAGATATGTGACGGGGCTAG	1570		
QY	1817	TCTTGAGTTTACTACCGTTATGTTTGGACGGCGCATGATGTATACATGATTTGAATGGGG	1876		
Db	1571	AATGGAGCGCGCTCGCTCGATACATATATAGGTACATTGGTGCACACTATTTGAAA	16334		
QY	1877	ACCGTCCGATGG 1888			
Db	1631	ATTGCTTATGG 1642			

RESULT 13
US-09-091-432-3
; Sequence 3, Application US/090914322
; Patent No. 5981837
; GENERAL INFORMATION:
; APPLICANT: Chapple, Clint

```

1 TITLE OF INVENTION: A Method For Regulation Of Plant Lignin Composition
2
3 FILE REFERENCE: 7024-325
4
5 CURRENT APPLICATION NUMBER: US/09/091,432
6
7 EARLIER FILING DATE: 1998-06-18
8
9 EARLIER APPLICATION NUMBER: PCT/US96/20094
10
11 EARLIER FILING DATE: 1996-12-19
12
13 EARLIER APPLICATION NUMBER: US 60/009,119
14
15 EARLIER FILING DATE: 1995-12-22
16
17 EARLIER APPLICATION NUMBER: US 60/013,388
18
19 EARLIER FILING DATE: 1996-03-14
20
21 NUMBER OF SEQ ID NOS: 3
22
23 SOFTWARE: Microsoft Word 2.0C
24
25 SEQ ID NO 3
26
27 LENGTH: 5156
28
29 TYPE: DNA
30
31 ORGANISM: Arabidopsis thaliana
32
33 US-09-091-432-3
34
35 Query Match 4.1% Score 81.4; DB 2; Length 5156;
36 Best Local Similarity 49.6%; Pred. No. 17e-12;
37 Matches 276; Conservative 0; Mismatches 266; Indels 15; Gaps 2;

```

QY	1334	TAAATTTTGTAGTTATATATTAACAGAGAGATGATATTTCAAGAGAAACAGACAGTTGGCGGT	1333
Db	3877	TGATTTTTTTTATTGATTTTATTAGACGTTATCTTTGGAGGAAACGAAACGTTAGCGTC	3936
QY	1394	CTTATTCGAGTGGATTCCTCGCTAGAGTGTCTTCATTCAGATATGCAATCAAGGTACA	1453
Db	3937	GGCGATTAGAGGGCCCTTTAACGAGATTATTCGAGGCCCCGAGATCTAAACGGGTCCA	3996
QY	1454	AAACGAGCTGAGTCAAGTAGTCGGGAAATCAAGAGCCCTAGATGAAATCTGACTTGGCTTC	1513
Db	3997	ACAAGAACTGGCCGAGTGGTTGAGCTTGAACAGAGATTGAAGAAATCCGACATCGAGAA	4056
QY	1514	ACTTCCATATCTTAACGGCTGTGGTGAAGAAAGTATGAGGTTTCACTCCAGGGCCACT	1573
Db	4057	GTTGACTTATCTCAATGACACATCTAAAGAAACCTTAAGATGACCCACCGATTCCTCT	4116
QY	1574	TCATATCATGGCCCGTTTGGCCATACAGACAGCATGTTAGTGTCTGTTTCCGSC	1633
Db	4117	CCT-----CCTCACGAAACCGGGAGAGACATAGTATCGACGGTTTCTCATTTCCAA	4170
QY	1634	AGGACACACACAGATGTGAAACATGTGGGCCGTATCGATATACACACGTGGGTGA	1699
Db	4171	GAATTCGTGTGATGATCAACGGCTTGGCCATAGAGACGACCCACCTTTGGACTGA	4233
QY	1694	TCCTTTGAGTTTAAACCTGAGAGGTTCTGTGGCAAAAGAGGTGAGGTTCGCT	1753
Db	4231	CCCGGACACGTTTAGACCATCGAGGTTTGTGAACCGGCGTACCGG-----ATTT	4281
QY	1754	TCTTGGTTCGGAATTGAAACTTGCACCTTGGGGTGGGTGGTGGATTTGGCCCGGGAA	1813
Db	4282	CAAAAGGAGCAATTTTCAGTTTATACGTTTACCGTTCGGGTGGTGAATGTGCCCCGGTAT	4341
QY	1814	GAATCTTGTTTACTACCGTTATGTTTGGACGGCGATGATGTTACATGAGTTTGAATG	1873
Db	4342	GCAATTAAGGTTATACGGCGTTGACTTAGCGGTGCATATATTAATACATGCTTCACGCTG	4401
QY	1874	GGAGCCGTCGATGTGA	1890
Db	4402	GAATTTACCTATGGGA	4418

```

/ RESULT 14
/ US-09-387-663-3
/ Sequence 3, Application US/09387663
/ Patent No. 6489538
/ GENERAL INFORMATION
/ APPLICANT: Chapelle, Clint
/ TITLE OF INVENTION: A Method For Regulation Of Plant Lignin Composition
/ FILE REFERENCE: 7024-325
/ CURRENT APPLICATION NUMBER: US/09/387,663
/ CURRENT FILING DATE: 1999-08-31

```

PRIOR APPLICATION NUMBER: US 09/091,432
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/US96/20094
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: US 60/009,119
PRIOR FILING DATE: 1995-12-22
PRIOR APPLICATION NUMBER: US 60/013,388
PRIOR FILING DATE: 1996-03-14
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Microsoft Word 2.0C
SEQ ID NO: 3
LENGTH: 5156
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-387-663-3

Query Match 4.1%; Score 81.4; DB 4; Length 5156;
Best Local Similarity 49.6%; Pred. No. 1.7e-12;
Matches 276; Conservative 0; Mismatches 266; Indels 15; Gaps 2;

QY 1334 TAAATTTTGAAGTATTTATACAGAGATGATATTCAGAGAACAGACAGTTGGGT 1393
DB 3877 TGAATTTTGAAGTATTTATACAGAGATGATATTCAGAGAACAGACAGTTGGGT 3936
QY 1394 CTTATCGAGTGAATCTCGTAGAGATGCTCTTATCAGATATGCAATCAAGGTACA 1453
DB 3937 GGGCATAGAGTGGCTTAAAGAGATTTATACAGAGATTTAAAGAGATTTAAAGAGTCCA 3996
QY 1454 AAGAGAGTGAATCAAGTATGAGAGAAATCAAGAGCCCTAGATATGATGAGTTGGGT 1513
DB 3997 ACAAAGAGTGGCTTAAAGAGATTTATACAGAGATTTAAAGAGATTTAAAGAGTCCA 4056
QY 1514 ACTTCATATCTAAGAGTATGAGAGAAAGATATGAGTTGAGTTGAGTTGAGTTGAGTTG 1573
DB 4057 GTTGAATTTATCAAGTATGAGAGAAATCAAGAGATTTAAAGAGATTTAAAGAGTCCA 4116
QY 1574 TTTATCAGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 1633
DB 4117 CCTTCT 4170
QY 1634 AAGGAGCAGAGAGTGAATGAGATGAGAGTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 1693
DB 4171 GAAATTCCTGATGATATCAAGCCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 4230
QY 1694 TCTTTGAGTTTAAAGTGAAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAGAGTCTGAG 1753
DB 4231 CCGGACAGTGTAGACATGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 4281
QY 1754 TCTTGGGTGGAGTTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1813
DB 4282 CAAAGGAGCAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 4341
QY 1814 GAATCTGTTTATCAAGTATGATTTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1873
DB 4342 GCAATAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 4401
QY 1874 GGGACGTCGATGTA 1890
DB 4402 GAAATTAACGATGGA 4418

RESULT 15

US-09-033-055A-1
Sequence 1, Application US/09033055A

PATENT INFORMATION:
GENERAL INFORMATION:
APPLICANT: OHKAWA, HIROKO
TITLE OF INVENTION: NOVEL CYTOCHROME P450 GENE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO
STREET: 1100 NEW YORK AVENUE, N.W.

CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,055A
FILING DATE: MARCH 2, 1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PAUL E. WHITE, JR.
REGISTRATION NUMBER: 32,011
REFERENCE/DOCKET NUMBER: 9437/251563
TELEPHONE: (202) 861-3651
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1722
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-033-055A-1

Query Match 3.9%; Score 79.2; DB 3; Length 1722;
Best Local Similarity 49.5%; Pred. No. 4.2e-12;
Matches 275; Conservative 0; Mismatches 263; Indels 18; Gaps 2;

QY 1319 AAAAGTAAAGTAAATTTTGAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1378
DB 849 ACATGCAATTTAGTAAATCAATATAGAGGCTTGGAGACCTTTTATTTATTTATTTATTT 908
QY 1379 AGACAGAGTGGGCTTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 1438
DB 909 TGAATTAAGTTCAAAATCAATTTAGATGAGGCAATGTTAGACATAGAGAAACATATACAT 968
QY 1439 GCAATCAACGATCAAAAGAGAGTGAATGATGAGTGAATGAGTGAATGAGTGAATGAGTGA 1498
DB 969 AATGAAAAAGTAAAGAGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGA 1028
QY 1499 ATCTGAGTGGCTTCACTTCCATATCTAAAGGCTGAGTGAAGAGTGAAGTGAAGTGAAGTGA 1558
DB 1029 AGCTGATATTTGAGCTCTCCATTTTGAAGTGAATTTGAAGTGAATTTGAAGTGAATTTGA 1088
QY 1559 TCTTCAAGGCTTCTATCATGAGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 1618
DB 1089 CCGACAGGCTCTCTTTTAAATTTGAGTGAATTTGAGTGAATTTGAGTGAATTTGAGTGAATTT 1142
QY 1619 TCGTCTTGTTCGGAGAGGACAGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATG 1678
DB 1143 CTATTTTGTCCAAAGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATG 1202
QY 1679 ACAGTGTGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAATG 1738
DB 1203 AACTATATGAGAGATCTTTTGAATTTAGCCGAAAGTGTGAGGCTTCAAAATTTGAGTGAAT 1262
QY 1739 GGTGAGTGTTCGTTCTTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTT 1798
DB 1263 C-----GTTGAGAGAGTGTGAGTGAATTTGAGTGAATTTGAGTGAATTTGAGTGAATTTG 1310
QY 1799 GATTTGCCCGGAGAGATCTTTGTTTATCTACCGTTATGTTTGAAGGCGAGTATGTT 1858
DB 1311 AATTTGCCCGGAGTATCTTTGCAACAGAGATTTACGCAATGATGTTGATGTTTATGTT 1370
QY 1859 ACATGAGTTGATG 1874
DB 1371 GAATTCATTTGATGG 1386

Tue Jan 20 17:43:17 2004

us-10-022-025a-1.rnt

Page 11

Search completed: January 16, 2004, 13:43:40
Job time : 129 secs

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rmpb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 03:16:41 ; Search time 538 Seconds
(without alignments)
10436.706 Million cell updates/sec

Title: US-10-022-025a-2

Perfect score: 1593

Sequence: 1 atggctacgaactcgaag.....tgcgcgttagcgagcttaa 1593

Scoring table:

Gapop 10.0, Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PC9_NEM_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEM_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/PC9_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_NEM_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US10_NEM_PUB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10_NEM_PUB.seq:*
17: /cgn2_6/prodata/1/pubpna/US60_NEM_PUB.seq:*
18: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1593	100.0	1593	US-09-349-385-5	Sequence 5, Appl
2	1593	100.0	1593	US-10-022-025a-2	Sequence 2, Appl
3	1360	85.4	2009	US-10-022-025a-1	Sequence 1, Appl
4	976	61.3	1902	US-09-349-385-2	Sequence 2, Appl
5	794	49.8	7869	US-09-349-385-3	Sequence 3, Appl
6	651.6	40.9	1611	US-09-349-385-7	Sequence 7, Appl
7	599.6	37.6	1764	US-10-163-198-20	Sequence 20, Appl
8	585.8	36.8	1905	US-10-163-198-24	Sequence 24, Appl
9	355	22.3	1799	US-09-349-385-13	Sequence 13, Appl
10	352	22.1	1930	US-10-163-198-34	Sequence 34, Appl
11	348.2	21.9	1990	US-09-349-385-9	Sequence 9, Appl
12	316.6	19.9	1934	US-10-163-198-22	Sequence 22, Appl
13	290.6	18.2	1687	US-10-163-198-40	Sequence 40, Appl
14	286	18.0	1545	US-10-163-198-96	Sequence 96, Appl
15	283	17.8	1758	US-10-163-198-94	Sequence 94, Appl

16	278.2	17.5	1597	15	US-10-163-198-16	Sequence 16, Appl
17	275.2	17.3	1585	15	US-10-163-198-92	Sequence 92, Appl
18	274.4	17.2	1539	15	US-10-163-198-18	Sequence 18, Appl
19	265.8	16.7	2924	15	US-10-163-198-26	Sequence 26, Appl
20	237.4	14.9	2087	10	US-09-349-385-11	Sequence 11, Appl
21	234	14.7	1146	15	US-10-163-198-12	Sequence 12, Appl
22	233.4	14.7	1578	15	US-10-163-198-1	Sequence 1, Appl
23	213.4	13.4	1121	15	US-10-163-198-8	Sequence 8, Appl
24	194.8	12.2	1357	15	US-10-163-198-36	Sequence 36, Appl
25	189.2	11.9	884	15	US-10-163-198-3	Sequence 3, Appl
26	156.4	9.8	442	15	US-10-163-198-38	Sequence 38, Appl
27	155.6	9.8	1720	15	US-10-163-198-3	Sequence 3, Appl
28	129.6	8.1	1708	9	US-09-796-556A-1	Sequence 1, Appl
29	128.2	8.1	1721	12	US-10-429-949-19	Sequence 19, Appl
30	126	7.9	778	15	US-10-163-198-14	Sequence 14, Appl
31	119.4	7.5	1527	9	US-09-331-267-1	Sequence 1, Appl
32	119.4	7.5	1549	9	US-09-331-267-3	Sequence 3, Appl
33	117.8	7.4	1527	9	US-09-331-267-2	Sequence 2, Appl
34	111.2	7.0	1673	12	US-10-431-273-42	Sequence 42, Appl
35	111.2	7.0	1714	12	US-10-431-273-43	Sequence 43, Appl
36	106.8	6.7	1770	12	US-10-411-115-7	Sequence 7, Appl
37	106.8	6.7	1933	14	US-10-067-534-1	Sequence 1, Appl
38	105.6	6.6	457	15	US-10-163-198-30	Sequence 30, Appl
39	104.8	6.6	1747	12	US-10-431-273-40	Sequence 40, Appl
40	104.8	6.6	1763	12	US-10-431-273-41	Sequence 41, Appl
41	101.6	6.4	1557	10	US-09-938-842A-2669	Sequence 2669, Ap
42	99.6	6.3	1545	13	US-10-356-153-55	Sequence 55, Appl
43	99.6	6.3	1545	15	US-10-142-231-55	Sequence 55, Appl
44	98.8	6.2	528	15	US-10-163-198-28	Sequence 28, Appl
45	97.2	6.1	1764	10	US-09-947-027-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1	
US-09-349-385-5	
Sequence 5, Application US/09349385	
Patent No. US20020152495A1	
GENERAL INFORMATION:	
APPLICANT: Ito, Toshio	
APPLICANT: Meyersowitz, Elliot	
TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT	
FILE REFERENCE: NBI-0002	
CURRENT APPLICATION NUMBER: US/09/349,385	
CURRENT FILING DATE: 1999-07-09	
EARLIER APPLICATION NUMBER: 60/115,967	
EARLIER FILING DATE: 1999-01-15	
NUMBER OF SEQ ID NOS: 13	
SOFTWARE: PatentIn Ver. 2.0	
SEQ ID NO 5	
LENGTH: 1593	
TYPE: DNA	
ORGANISM: Arabidopsis thaliana	
FEATURE:	
OTHER INFORMATION: CDNA	
US-09-349-385-5	
QY	1 atggctacgaactcgaagcttgaatcttggccctttgccaatgacgagcttca 60
DB	1 atggctacgaactcgaagcttgaatcttggccctttgccaatgacgagcttca 60
QY	61 agccaaacaaacttgccttctccctctgcgcgtgacatcatgtgctgcacatct 120
DB	61 agccaaacaaacttgccttctccctctgcgcgtgacatcatgtgctgcacatct 120
QY	121 ctctcttattgacatcattccgctgacatctgctgggggaattactcttggcggggtta 180

Db 121 CTCTTCTATGAGACCTATCCCGGTGAGACTGCTTGAGGAAATACCTCTTCGCGCGTTA 180
 QY 181 ATATCCGGTTCATCAAAACCGGAAAGTTATTCGCGGTCCAAAGGCTCCCTTGGTT 240
 Db 181 ATATCCGGTTCATCAAAACCGGAAAGTTATTCGCGGTCCAAAGGCTCCCTTGGTT 240
 QY 241 GGAAGCATGTCATCATGTCAAGCACTCTAGCTCAGCAGAAATGCTGATGCACTGAG 300
 Db 241 GGAAGCATGTCATCATGTCAAGCACTCTAGCTCAGCAGAAATGCTGATGCACTGAG 300
 QY 301 AAATTCGAGCCAAAGGCTCATGCTTTCAGCTTAAAGAGACTCCGCTGATGCTACG 360
 Db 301 AAATTCGAGCCAAAGGCTCATGCTTTCAGCTTAAAGAGACTCCGCTGATGCTACG 360
 QY 361 TGCAATCCCGACGTAGCGAAAGAGATTCTGAATAGCCCGGTTTTCGCTGATGCAACG 420
 Db 361 TGCAATCCCGACGTAGCGAAAGAGATTCTGAATAGCCCGGTTTTCGCTGATGCAACG 420
 QY 421 AAAGATGCGCTTACTCATCTGATGTTTAAAGAGCAATGGTTTTCACACACACG 480
 Db 421 AAAGATGCGCTTACTCATCTGATGTTTAAAGAGCAATGGTTTTCACACACACG 480
 QY 481 TACTGCGAAGCGCTTCGCGGTATGCTGCAACCATCTCTTATGATCAAAACCAATCAGA 540
 Db 481 TACTGCGAAGCGCTTCGCGGTATGCTGCAACCATCTCTTATGATCAAAACCAATCAGA 540
 QY 541 AGAGCCGAGACGCAACGAGAGTATCTCAACCAAGATGTTGAGTTTCTTGAACACAG 600
 Db 541 AGAGCCGAGACGCAACGAGAGTATCTCAACCAAGATGTTGAGTTTCTTGAACACAG 600
 QY 601 AGTGTATCGAAACCTGTTTGTTCGAGAGTTGCTTAAACGCGGTGCTTAAACAATG 660
 Db 601 AGTGTATCGAAACCTGTTTGTTCGAGAGTTGCTTAAACGCGGTGCTTAAACAATG 660
 QY 661 ATGTGCTCTGTATTCGCAACAGAGTATGAGCTTGAACCAATGTTGAGTTTCTG 720
 Db 661 ATGTGCTCTGTATTCGCAACAGAGTATGAGCTTGAACCAATGTTGAGTTTCTG 720
 QY 721 ATGTGCGAAGAGTTATGATGTTGCTCGAAGCTGATGAGTCACTGATCACTTCCT 780
 Db 721 ATGTGCGAAGAGTTATGATGTTGCTCGAAGCTGATGAGTCACTGATCACTTCCT 780
 QY 781 CTATCGAGATTGATCTCAAGACTCCGCTCTAGATGTTTCAACATCTGATCAACAG 840
 Db 781 CTATCGAGATTGATCTCAAGACTCCGCTCTAGATGTTTCAACATCTGATCAACAG 840
 QY 841 AACCGGTTTGTATCCCGGATTAATCCGCAACACCGTATCAAAACCGGATGTTT 900
 Db 841 AACCGGTTTGTATCCCGGATTAATCCGCAACACCGTATCAAAACCGGATGTTT 900
 QY 901 GATTTCGTGAGCGTTTGTCTCTCCCTCATGAGTTTCAATTAATTAATCCGACCG 960
 Db 901 GATTTCGTGAGCGTTTGTCTCTCCCTCATGAGTTTCAATTAATTAATCCGACCG 960
 QY 961 ATGCGCGTTCTTTGGAGATGATTAATCAAGAGACAGACAGATTGCGGCTTAT 1020
 Db 961 ATGCGCGTTCTTTGGAGATGATTAATCAAGAGACAGACAGATTGCGGCTTAT 1020
 QY 1021 TGGATCTCTGAGAGTGGTCTTCTCATCCAGATATGATCAACGATCAAAAGAG 1080
 Db 1021 TGGATCTCTGAGAGTGGTCTTCTCATCCAGATATGATCAACGATCAAAAGAG 1080
 QY 1081 GATCAAGTATGTCGGGAATCAAGAGCCTAGATGATCTGACTTGGCTTCACTTCAT 1140
 Db 1081 GATCAAGTATGTCGGGAATCAAGAGCCTAGATGATCTGACTTGGCTTCACTTCAT 1140
 QY 1141 CTAAAGCGCTGTGTGAAGAGATTAATGAGGCTTCACTTCAGAGCCCACTTATCAT 1200
 Db 1141 CTAAAGCGCTGTGTGAAGAGATTAATGAGGCTTCACTTCAGAGCCCACTTATCAT 1200
 QY 1201 GCCCGTTTGGCCATTAACAGACAGATGTTGATGCTGTTTTCGCGACAGGACACA 1260
 Db 1201 GCCCGTTTGGCCATTAACAGACAGATGTTGATGCTGTTTTCGCGACAGGACACA 1260

QY 1261 GCATGTGTAACATGTGGCCGTATGCGATGATCCACAGTGTGGTTGATCTTTGGAG 1320
 Db 1261 GCATGTGTAACATGTGGCCGTATGCGATGATCCACAGTGTGGTTGATCTTTGGAG 1320
 QY 1321 TTTAACTGAGAGGTTTGTGGCAAAAGAGTGAAGTGAAGTTTGTGGTTTGGT 1380
 Db 1321 TTTAACTGAGAGGTTTGTGGCAAAAGAGTGAAGTGAAGTTTGTGGTTTGGT 1380
 QY 1381 GATTGAGACTTGACCTTTCCGGTCCGCTGATGATTTTCCCGGGAAGATCTTGGT 1440
 Db 1381 GATTGAGACTTGACCTTTCCGGTCCGCTGATGATTTTCCCGGGAAGATCTTGGT 1440
 QY 1441 TTTACTACCGTTATGTTTTCGAGCGGATGATGATTTTCAATGAGTTTGAATGGGAC 1500
 Db 1441 TTTACTACCGTTATGTTTTCGAGCGGATGATGATTTTCAATGAGTTTGAATGGGAC 1500
 QY 1501 GATGTAACGCGCTTGACTTATCTGAGAACTGAGGCTTTCCTTGGAGATGCTATCT 1560
 Db 1501 GATGTAACGCGCTTGACTTATCTGAGAACTGAGGCTTTCCTTGGAGATGCTATCT 1560
 QY 1561 CTTCCTGCTAAATTCGCGCGTAAAGCGGAGTTAA 1593
 Db 1561 CTTCCTGCTAAATTCGCGCGTAAAGCGGAGTTAA 1593
 RESULT 2
 US-10-022-025A-2
 ; Sequence 2, Application US/10022025A
 ; Publication No. US20030092014A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Coleman, John R.
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Polypeptides for Catabolism of Absc.
 ; FILE REFERENCE: 3310 0003
 ; CURRENT APPLICATION NUMBER: US/10/022,025A
 ; PRIOR FILING DATE: 2003-01-09
 ; PRIOR FILING DATE: 2000-12-13
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 1593
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1593)
 US-10-022-025A-2
 Query Match 100.0%; Score 1593; DB 15; Length 1593;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1593; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCTAACGAACCTGAAAGCTCTTAATCTTGGCCCTTTGTCCAATGAGCGTTCTA 60
 Db 1 ATGGCTAACGAACCTGAAAGCTCTTAATCTTGGCCCTTTGTCCAATGAGCGTTCTA 60
 QY 61 AGCCAAACCAACCTTCTCTTCCCTCTGCGCGTCAACATCATCTGCTCCCATATCT 120
 Db 61 AGCCAAACCAACCTTCTCTTCCCTCTGCGCGTCAACATCATCTGCTCCCATATCT 120
 QY 121 CTCTTTTATGAGCTTATCCCGTGTGACCTGCTGGGGAATPAACTCTTGGCGGTTA 180
 Db 121 CTCTTTTATGAGCTTATCCCGTGTGACCTGCTGGGGAATPAACTCTTGGCGGTTA 180
 QY 181 ATATCCGTTTCAATCAAAACCGGAAAGCTTATCCCGGTCAAAAGGCTTCCCTTGGTT 240
 Db 181 ATATCCGTTTCAATCAAAACCGGAAAGCTTATCCCGGTCAAAAGGCTTCCCTTGGTT 240
 QY 241 GGAAGCATGTCATCATGTGCAAGCACTCTAGCTCAGCAGAAATGCTGATGCACTGAG 300
 Db 241 GGAAGCATGTCATCATGTGCAAGCACTCTAGCTCAGCAGAAATGCTGATGCACTGAG 300

QY 301 AAATCGGAGCCAGAGAGGCTCATGGCTTTGAGAGAGATCGCGGATCGTCAG 360
 Db 301 AAATCGGAGCCAGAGAGGCTCATGGCTTTGAGAGAGATCGCGGATCGTCAG 360
 QY 361 TGCATATCCGACGTAGGAGAAAGATTTGAAATAGCCCGGCTTTTGTGATGACCGGTT 420
 Db 361 TGCATATCCGACGTAGGAGAAAGATTTGAAATAGCCCGGCTTTTGTGATGACCGGTT 420
 QY 421 AAAGAAATCGGCTTACTCATGATGATTTAAACAGAGCAATGGTTTGCACACACGAGTGT 480
 Db 421 AAAGAAATCGGCTTACTCATGATGATTTAAACAGAGCAATGGTTTGCACACACGAGTGT 480
 QY 481 TACTGGGAAACGCTTGCCTGATCGCTTGCACCAATCTCTTTAGTCAAAAACAAATCAGA 540
 Db 481 TACTGGGAAACGCTTGCCTGATCGCTTGCACCAATCTCTTTAGTCAAAAACAAATCAGA 540
 QY 541 AGAGCCGAGACGCAACGAGATGATCTCAAGCCAGATGGTTGAGTTCTTGAAAAACAG 600
 Db 541 AGAGCCGAGACGCAACGAGATGATCTCAAGCCAGATGGTTGAGTTCTTGAAAAACAG 600
 QY 601 AGTAGTAACGAACCTGTTTTGTTGCTGAGATTGCTTAAAAACGCGCTGCTTAAACACATG 660
 Db 601 AGTAGTAACGAACCTGTTTTGTTGCTGAGATTGCTTAAAAACGCGCTGCTTAAACACATG 660
 QY 661 ATGTGCTCTGATTTCCGACAGAGTATGAGCTTGAAAAACCATGTGAGTTAGCTGA 720
 Db 661 ATGTGCTCTGATTTCCGACAGAGTATGAGCTTGAAAAACCATGTGAGTTAGCTGA 720
 QY 721 ATGTGCTCTGATTTCCGACAGAGTATGAGCTTGAAAAACCATGTGAGTTAGCTGA 780
 Db 721 ATGTGCTCTGATTTCCGACAGAGTATGAGCTTGAAAAACCATGTGAGTTAGCTGA 780
 QY 781 CTATCGAGTTGATCTCTCAAGAATCTCGGCTTAAATGTTCCACTGCTGACCAAGTA 840
 Db 781 CTATCGAGTTGATCTCTCAAGAATCTCGGCTTAAATGTTCCACTGCTGACCAAGTA 840
 QY 841 AACCGGTTGATCCCGGATTAATATCGGAACACCGTATCAACCGGTGATTTGCTCGT 900
 Db 841 AACCGGTTGATCCCGGATTAATATCGGAACACCGTATCAACCGGTGATTTGCTCGT 900
 QY 901 GATTTGCTGACGTTTTGCTCTCCCTCCATGATGATTAATTAATCCGACCCGACATA 960
 Db 901 GATTTGCTGACGTTTTGCTCTCCCTCCATGATGATTAATTAATCCGACCCGACATA 960
 QY 961 ATGCGCGTTCTTTGGGAGATGATATTAGAGAAACAGACACAGTTGCGGCTTAAATCGAG 1020
 Db 961 ATGCGCGTTCTTTGGGAGATGATATTAGAGAAACAGACACAGTTGCGGCTTAAATCGAG 1020
 QY 1021 TGAATCTCTGCTAGAGATGATCTTCATCCAGATATGCAATCAACGGTACAAACGAGCTG 1080
 Db 1021 TGAATCTCTGCTAGAGATGATCTTCATCCAGATATGCAATCAACGGTACAAACGAGCTG 1080
 QY 1081 GATCAATAGTATCGGGAATCAAGAGCCCTAGATGAATCTTGAGCTTCACTTCATAT 1140
 Db 1081 GATCAATAGTATCGGGAATCAAGAGCCCTAGATGAATCTTGAGCTTCACTTCATAT 1140
 QY 1141 CTAAACGCTGTGTAAGAAAGATTAAGAGCTTCATCTCCAGAGCCACATTCATATATG 1200
 Db 1141 CTAAACGCTGTGTAAGAAAGATTAAGAGCTTCATCTCCAGAGCCACATTCATATATG 1200
 QY 1201 GCCCGTTTGGCATTAACAGACAGATGTTGATGATGTTGTTGTTCCGGAGGAGACACA 1260
 Db 1201 GCCCGTTTGGCATTAACAGACAGATGTTGATGATGTTGTTGTTCCGGAGGAGACACA 1260
 QY 1261 GCATATGTAACATGTGGGCGGATTCGATGATCAACAGTGGGTGATCTTTGGAG 1320
 Db 1261 GCATATGTAACATGTGGGCGGATTCGATGATCAACAGTGGGTGATCTTTGGAG 1320
 QY 1321 TTTAAACCTGAGAGGTTGTTGCAAAAGAGTGAAGTGGATTTTCGTTCTTGGGCTG 1380
 Db 1321 TTTAAACCTGAGAGGTTGTTGCAAAAGAGTGAAGTGGATTTTCGTTCTTGGGCTG 1380

QY 1381 GATTGAGACCTTGACCTTTTCGGGTCGGTCTGCGATTTGCCCCGGGAGAAATCTTGGT 1440
 Db 1381 GATTGAGACCTTGACCTTTTCGGGTCGGTCTGCGATTTGCCCCGGGAGAAATCTTGGT 1440
 QY 1441 TTTACTACCTTATGTTTGAACGCGCATGATGTTAATGATGTTGAATGCGGACCGTCC 1500
 Db 1441 TTTACTACCTTATGTTTGAACGCGCATGATGTTAATGATGTTGAATGCGGACCGTCC 1500
 QY 1501 GATGCTAACGCGCTGACTTATCTGAGAACTGAGAGCTTTCTTCCGAGATGCTAATCT 1560
 Db 1501 GATGCTAACGCGCTGACTTATCTGAGAACTGAGAGCTTTCTTCCGAGATGCTAATCT 1560
 QY 1561 CTTCCTGCTAATTTGCGCGCTGAGGCGCACTTAA 1593
 Db 1561 CTTCCTGCTAATTTGCGCGCTGAGGCGCACTTAA 1593
 RESULT 3
 US-10-022-025a-1
 ; Sequence 1, Application US/10022025A
 ; Publication No. US2003092014A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Coleman, John R.
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Polypeptides for Catabolism of Absci
 ; TITLE OF INVENTION: Acid
 ; FILE REFERENCE: 3310 0003
 ; CURRENT APPLICATION NUMBER: US/10/022,025A
 ; PRIOR FILING DATE: 2003-01-09
 ; PRIOR FILING DATE: 2000-12-13
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 2009
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: gene
 ; LOCATION: (1)..(2009)
 US-10-022-025a-1
 Query Match 85.4%; Score 1360; DB 15; Length 2009;
 Best Local Similarity 87.7%; Pred. No. 0;
 Matches 1593; Conservative 0; Mismatches 0; Indels 223; Gaps 1;

QY 1 ATGGCTAGAAACCTGAAAGCTCCTTAATCTTGGCCCTTTTGTCAATGAGAGGTTCTA 60
 Db 1 ATGGCTAGAAACCTGAAAGCTCCTTAATCTTGGCCCTTTTGTCAATGAGAGGTTCTA 60
 QY 61 AGCCAAACCAACCTTGCCTTCTCCCTCTCGCGGTCAATCATCTGAGTCCCATATCT 120
 Db 61 AGCCAAACCAACCTTGCCTTCTCCCTCTCGCGGTCAATCATCTGAGTCCCATATCT 120
 QY 121 CTCTCTTATGACCTTATCCCGGTGAGACCTGCTTGGGGGAAATCTCTTCCGCGGTTA 180
 Db 121 CTCTCTTATGACCTTATCCCGGTGAGACCTGCTTGGGGGAAATCTCTTCCGCGGTTA 180
 QY 281 CTCTCTTATGACCTTATCCCGGTGAGACCTGCTTGGGGGAAATCTCTTCCGCGGTTA 340
 Db 281 CTCTCTTATGACCTTATCCCGGTGAGACCTGCTTGGGGGAAATCTCTTCCGCGGTTA 340
 QY 181 ATATCCGCTTATCAAAAACCGGAAAGTTATCCCGGTCCAAAAGCTTCCCTTGGT 240
 Db 181 ATATCCGCTTATCAAAAACCGGAAAGTTATCCCGGTCCAAAAGCTTCCCTTGGT 240
 QY 341 ATATCCGCTTATCAAAAACCGGAAAGTTATCCCGGTCCAAAAGCTTCCCTTGGT 400
 Db 341 ATATCCGCTTATCAAAAACCGGAAAGTTATCCCGGTCCAAAAGCTTCCCTTGGT 400
 QY 241 GGAAGCATGTCACTCATGTCAAGACATCTTACTCAACGAGCAATCCGTGATGACGTGAG 300
 Db 241 GGAAGCATGTCACTCATGTCAAGACATCTTACTCAACGAGCAATCCGTGATGACGTGAG 300
 QY 401 GGAAGCATGTCACTCATGTCAAGACATCTTACTCAACGAGCAATCCGTGATGACGTGAG 460
 Db 401 GGAAGCATGTCACTCATGTCAAGACATCTTACTCAACGAGCAATCCGTGATGACGTGAG 460
 QY 301 AAATCGAGGCAAGAGGCTCATGAGCTTTCAGTTAGAGAGACTCGGATGATGATCAG 360
 Db 301 AAATCGAGGCAAGAGGCTCATGAGCTTTCAGTTAGAGAGACTCGGATGATGATCAG 360
 QY 461 AAATCGAGGCAAGAGGCTCATGAGCTTTCAGTTAGAGAGACTCGGATGATGATCAG 520
 Db 461 AAATCGAGGCAAGAGGCTCATGAGCTTTCAGTTAGAGAGACTCGGATGATGATCAG 520
 QY 361 TGCATATCCGACGTAGGAGAAAGATTTGAAATAGCCCGGCTTTTGTGATGACCGGTT 420
 Db 361 TGCATATCCGACGTAGGAGAAAGATTTGAAATAGCCCGGCTTTTGTGATGACCGGTT 420
 QY 420 TGCATATCCGACGTAGGAGAAAGATTTGAAATAGCCCGGCTTTTGTGATGACCGGTT 480
 Db 420 TGCATATCCGACGTAGGAGAAAGATTTGAAATAGCCCGGCTTTTGTGATGACCGGTT 480

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rnpb

Page 4

QY	42	AAAGAAATCGGCTTAACCTCACTGATAGTTTAAACAGAGAAATGGTATTGGACACACGGTGT	480
Db	581	AAAGAAATCGGCTTAACCTCACTGATAGTTTAAACAGAGAAATGGTATTGGACACACGGTGT	640
QY	481	TACTGGGAAACGCTTCGCGGTATGCTTCGAAACCATCTCTTTAGTACAAACCAATCAGA	540
Db	641	TACTGGGAAACGCTTCGCGGTATGCTTCGAAACCATCTCTTTAGTACAAACCAATCAGA	700
QY	541	AGAGCCGAGACGCAAGACGAGTGAATCTCAAGCCGAGATGGTGAAGTTCTTGAAGAAACAG	600
Db	701	AGAGCCGAGACGCAAGACGAGTGAATCTCAAGCCGAGATGGTGAAGTTCTTGAAGAAACAG	760
QY	601	AGTAGTAAAGCAACCTGCTTTTGTGTGTAGTTGCTTAAACGGCGTCCTTAAACAACATG	660
Db	761	AGTAGTAAAGCAACCTGCTTTTGTGTGTAGTTGCTTAAACGGCGTCCTTAAACAACATG	820
QY	661	ATGTCCTGTATTCGGAACAAGATGAGCTTGAAAAAACAATGTGATTAAGTGA	720
Db	821	ATGTCCTGTATTCGGAACAAGATGAGCTTGAAAAAACAATGTGATTAAGTGA	880
QY	721	ATGTCGGAAGAGTTATGATTTGCTCGGAACGTTGAATTTGAGATGATCACTTCCTTGG	780
Db	881	ATGTCGGAAGAGTTATGATTTGCTCGGAACGTTGAATTTGAGATGATCACTTCCTTGG	940
QY	781	CTATCGGAGTTTGATCTCTAAGAACTCCGAGCTAGATGTTCCACACTCGAACCAAGATA	840
Db	941	CTATCGGAGTTTGATCTCTAAGAACTCCGAGCTAGATGTTCCACACTCGAACCAAGATA	1000
QY	841	AACCGGTTTGATATCCCGGATTTATATCCGAAACCGGTAACTCAACCCGTGATTTGCTCGT	900
Db	1001	AACCGGTTTGATATCCCGGATTTATATCCGAAACCGGTAACTCAACCCGTGATTTGCTCGT	1060
QY	901	GATTTGCTGACGTTTTGCTCTCCCTCCATGAGTTCAATATAAATTATCCGACCGGACATA	960
Db	1061	GATTTGCTGACGTTTTGCTCTCCCTCCATGAGTTCAATATAAATTATCCGACCGGACATA	1120
QY	961	ATGCGCGTCTCTTG-----	974
Db	1121	ATGCGCGTCTCTTGCGGATGACACACCATTTATTGTAATTAATTCTTAATTATATTGT	1180
QY	975	-----	974
Db	1181	TGAAAAATGCTTAGATTAATTAGATTAACATGAATTTGAGACTCATGTGACGT	1240
QY	975	-----	974
Db	1241	TGTGGAATTAATTAAGCATTTAGAAAGTTTTTTGTTGACATCAAAATTAATTAATTTAGAT	1300
QY	975	-----GGA	977
Db	1301	TTTATTAACAGTTTCTATTAATAAGTAAAAAGTAAATTTTGAATTAATTATCAAGGA	1360
QY	978	GATGATTAATCAAGAGAAACAGACAGTTGGGGTCTTAATGAGTGAATCTCGTAAAGAT	1037
Db	1361	GATGATTAATCAAGAGAAACAGACAGTTGGGGTCTTAATGAGTGAATCTCGTAAAGAT	1422
QY	1038	GATCTTTCAATCGAGATATGCAATCAACGCTAACAAACGAGCTGATCAAGTATGTCGGAA	1097
Db	1421	GATCTTTCAATCGAGATATGCAATCAACGCTAACAAACGAGCTGATCAAGTATGTCGGAA	1480
QY	1098	ATCAAGAGCCCTAGATGAATCTGAATTTGAGCTTCACTTCAATCTAAACGGCTGTGTGA	1157
Db	1481	ATCAAGAGCCCTAGATGAATCTGAATTTGAGCTTCACTTCAATCTAAACGGCTGTGTGA	1540
QY	1158	AGAGATTTGAGGCTTCACTCTCCAGGCCCACTTCTATCATGGGCCGTTTGCCATAAC	1217
Db	1541	AGAGATTTGAGGCTTCACTCTCCAGGCCCACTTCTATCATGGGCCGTTTGCCATAAC	1600
QY	1218	AGACACGATGTTGATGTCGTCTTGTTCCGACAGGAGCAACAGCAATGGTGAACATGTCG	1277
Db	1601	AGACACGATGTTGATGTCGTCTTGTTCCGACAGGAGCAACAGCAATGGTGAACATGTCG	1666
QY	1278	GCGCGTATCCGATATCCACAGTGTGGGTGATCTTTGAGATTTAACTGAGAGTT	1337

Db	1661	GCGCGATCGCATGATCCACACGCTGTGGCTTGATCCTTTGGAGTTTAAACCTGGAGGTT	1720
Qy	1338	CGCGGAAAAGAGTGAGCGTGAGTTTCGGTCTTGGTCGGATTGAGCTTGAC	1337
Db	1721	CGGCGAAAAGAGTGAGTGAGCTTTTCGTTCTTGGATTCGAGATTGAGCTTGAC	1780
Qy	1398	TTTCGGGTGGGTCGTCGGAATTGGCCCGGAGAAACTTGGTTTACTACCGTTATGTT	1457
Db	1781	TTTCGGGTGGGTCGTCGGAATTGGCCCGGAGAAACTTGGTTTACTACCGTTATGTT	1840
Qy	1458	TTGAGCGGCGATGATGTATCATGATGTGAATGGGACCGTCCGATGTGAACGGCGTTGA	1517
Db	1841	TTGAGCGGCGATGATGTATCATGATGTGAATGGGACCGTCCGATGTGAACGGCGTTGA	1900
Qy	1518	CTTATCTGAAATCGAGGCTTTTTCGAGATGCGTAACTCCTTCGCTAAATTCG	1577
Db	1901	CTTATCTGAAATCGAGGCTTTTTCGAGATGCGTAACTCCTTCGCTAAATTCG	1960
Qy	1578	CGGTAGGCGCACTTAA	1593
Db	1961	CGGTAGGCGCACTTAA	1976

```

RESULT 4
US-09-349-385-2
: Sequence 2, Application US/09349385
: Patent No. US20020152495A1
: GENERAL INFORMATION:
: APPLICANT: Ico, Toshio
: APPLICANT: Fromm, Michael
: APPLICANT: Meyerowitz, Elliot
: TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT
: FILE REFERENCE: MB1-0002
: CURRENT APPLICATION NUMBER: US/09/349,385
: EARLIER FILING DATE: 1999-07-09
: EARLIER FILING DATE: 1999-01-15
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO: 2
: LENGTH: 1902
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: OTHER INFORMATION: cDNA
US-09-349-385-2

Query Match          61.3%; Score 976; DB 10; Length 1902;
Best Local Similarity 78.0%; Pred. No. 0;
Matches 128; Conservative 0; Mismatches 320; Indels 30; Gaps 4;

QY      19 AGCTCCTTAATCTTTGCCCTTTTGTCCTCAATGCAAGGCTTCTAAGCCAAACCAACCTTGGC 78
Db      90 AGTAGCTACTTTTGCCCTCTTGTCCTCAATGTAAGCTCTCTTACTCAACCAATCTTGGC 149

QY      79 TTCTCCCTCTGCGCGGTCACAATCATGTGGCTGCGCATATCTCTTTCTTAGACCTAT 138
Db      150 CTCTCTCTCTGTAAGCTCCCTAGCTTCTCTGCTCTTCTCTCTCTCTCTCTGCTGAT 209

QY      133 CCGGTGAAGCTGCTGGGGGAAATACCTTTTGCGCCGGTTAATCCGGTTCAACAA 198
Db      210 CCGGAGGACCCGGACGGGGAAGTAATTTCTCAACGC-----CGCGT 254

QY      199 ACCGGAACGTTATCCCGGTCCAAAGAGCTTCCCTTTGGTGAAGACATGCACTCATG 258
Db      225 CAACACACGTAATCCCGGGCCAAAGGCTTACCTTTGTGGGAAGCATGCTCTCATG 314

QY      259 TCAGAAGCTAAGCTCAACCGACGAATGCGTGAATGACGTGAAGAAATTCGAGGACAAAGG 318
Db      315 TCAGACCTGAGCTCAACCGTGTGATAGCCGCAACCGCAGAGAAATTTGAGCCGAAGG 374

QY      319 CTCATGGCTTCAGCTTAGAGAGAGACTGCGTATGTCATCGATGCAATCCGAGCTAGCG 378

```

```

Db      375  TTAATGGGCTTAAAGTTGGAGAACTGGCTGATGCTGATGCAATCTCGATGAGTACT 434
Qy      379  AAAGAGATTCTGAATAGCCCGGTTTTTCTGATGCAACGGGTTAAAGAAATCGGCTTACTCA 438
Db      435  AAAGAGATTCTAAACAGTCCGGTTTTTGGCTGACCCCGGTTAAAGAAATCAAGCTTATTC 494
Qy      439  CTGATGTTTAAACAGAGCAATTGGTTTGGACCAACGGGTTTACTGGCGAAGCGTTCGC 498
Db      495  CTCATGTTTAAACCGGCTAATCGGTTTCTCTTACGGGTTTACTGGCGAAGCGTTCGGA 554
Qy      499  CGTATCGCTTCAACCAATCTCTTATGTAACAAACAAATCAGAAAGCGGAGCGCAACA 558
Db      555  AAAATCGGCTTAAATCATCTTTTACGCCGAAACAGATTAACGTTCCGAAACGCAAGAGA 614
Qy      559  CGAGTATCTCAAGCCAGATGGTGAATTTCTTGAACAAACAGATTAAGCA---CCG 615
Db      615  ACGGTATCGCAATCAATCGTAAGTGTCTCAACAAACAGATTAACCAAGGCTTC 674
Qy      616  TGTATTGCTGATGCTTAAACCGGCTGCTTAAACAGATGATGCTGCTGATTC 675
Db      675  TGTTCGCACTGATCTTGAATCAAAACGGATCGCTTAAATACATGATGCTGCTGATTC 734
Qy      676  GGAAGAGATGAGCTTGAACAAACCATG-----TTGAGTTAGTGAATGCTC 726
Db      735  GGAAGAGATGAGCTTGAAGAGATGAAGAGATGAAGAGATGAGCTGATGAATGCTG 794
Qy      727  GAAGAGATGATGATTTGCTGCGAAGTGAATGATGATGATGATGATGATGATGATG 786
Db      795  GAGAGAGATGATGATTTACTCGGTACACGTAATGAGACCGATCATCTCCAGGCTCTC 854
Qy      787  GAGTTGATCTCAAAAGCTCGGCTGATGATGATGATGATGATGATGATGATGATG 846
Db      855  GAATTTGATCTCAAAAGATCCCGCTGATGATGATGATGATGATGATGATGATGATG 914
Qy      847  TTTGATCCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 906
Db      915  TTTGATCCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 974
Qy      907  GTGAGCTTTGCTCCCTCCATGATGATGATGATGATGATGATGATGATGATGATG 966
Db      975  GTTGAAGTATGCTCTCTCGATGATGATGATGATGATGATGATGATGATGATGATG 1034
Qy      967  GTTCTTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1026
Db      1035  GTTCTAATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1094
Qy      1027  CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1086
Db      1095  CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1154
Qy      1087  GTATCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1146
Db      1155  ATCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1214
Qy      1147  GCTGAGGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1206
Db      1215  GCTGAGGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1274
Qy      1207  TTGAGCATTAACAGACAGATGATGATGATGATGATGATGATGATGATGATGATG 1266
Db      1275  TTACCAATCAACAGACAGATGATGATGATGATGATGATGATGATGATGATGATG 1334
Qy      1267  GTGAACATGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1326
Db      1335  GTGAACATGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1394
Qy      1327  CCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1386
Db      1395  CCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1454
Qy      1387  AGACTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1446

```

```

Db      1455  AGGCTTGAACCGTTGCGGCTCGGCTCGGCTTGGCCCGGAGAAATCTTGGTTGACC 1514
Qy      1447  ACCGTTATGTTTGAACGCGCATGATGATGATGATGATGATGATGATGATGATGATG 1503
Db      1515  ACCGTTATGTTTGAACGCGCATGATGATGATGATGATGATGATGATGATGATGATG 1574
Qy      1504  GGTAAAGCGGCTTACTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1563
Db      1575  GAGAGAGCGGCTTACTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1634
Qy      1564  CTGCTTAATTTGCGCGTGAAGCGAGTT 1591
Db      1635  GCTGCTTAATTTAGCCCGAGCGGAGTT 1662

RESULT 5
US-09-349-385-3
; Sequence 3, Application US/09349385
; Patent No. US20020152495A1
; GENERAL INFORMATION:
; APPLICANT: Ito, Toshiro
; APPLICANT: Fromm, Michael
; APPLICANT: Meyerowitz, Elliot
; TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT
; FILE REFERENCE: MBI-0002
; CURRENT APPLICATION NUMBER: US/09/349,385
; CURRENT FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 60/115,967
; EARLIER FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 3
; LENGTH: 7869
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: genomic
US-09-349-385-3

Query Match      49.8%; Score 794; DB 10; Length 7869;
Beet Local Similarity 70.3%; Pred. No. 7,66-259;
Matches 1238; Conservative 0; Mismatches 320; Indels 202; Gaps 5;

```

```

Qy      19  AGCTCTTAATCTTTGCGCTTTGTCCAATGACGCTTCTAAGCCAAACCACTTGGC 78
Db      2154  AGTACTTACTTTGGCCCTCTTGCCAAATGATGATGATGATGATGATGATGATGATG 2213
Qy      79  TTCTCCCTCCCTGCGGCTGACATCATGCTGCTGCAATATCTCTTCTTATGACCTAT 138
Db      2214  CTCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2273
Qy      139  CCGGATGACCTGCTTGGGAGAAATATCCTCTTGGCGGCTTAATATCGGTTATACAAA 198
Db      2274  CCGGAGAGACCGGATGAGGAAAGATCTTCCACCG-----CCGCGCT 2318
Qy      199  ACCGAAACGTTATTCCTGCTTCAAAAGGCTTCCCTTGGTTGGAAGCATGCACTCATG 258
Db      2319  CAAACACCGTATATCCCGGCGCAAGAGGCTTACTTTTGGAGAGATGCTCTCATG 2378
Qy      259  TCAGAGATCTAGCTCAACGACGATCGCTGATGATGATGATGATGATGATGATGATG 318
Db      2375  TCMAACATCTGCTGATCAACGTTGATGATGATGATGATGATGATGATGATGATGATG 2438
Qy      319  CTCATGCTTTAAGCTTGAAGAGACTGCGGATGATGATGATGATGATGATGATGATG 378
Db      2439  TTAATGCGTTTAAAGTTGAGAAATCGCGATGATGATGATGATGATGATGATGATGATG 2498
Qy      379  AAAGAGATCTGAATAGCCCGGTTTTTCTGATGATGATGATGATGATGATGATGATG 438
Db      2499  AAAGAGATCTTAAACAGTCCGGTTTTTCTGATGATGATGATGATGATGATGATGATG 2558
Qy      439  CTGATGTTTAAACAGAGCAATGCTTTTGAACACAGGCTTATGATGATGATGATGATG 498

```

Db 2559 CTCATGTTAACGTCCTATCGGTTTCGCTTACGCGCTTACTGCGAACCCTTGA 2618
 QY 499 CGATCGCTTCGACCATCTCTTAACTAAGAAACAAATCAGAGAGCCGAGCAAGA 558
 Db 2619 AAATCGCGCTTATCATCTTTTCAAGCCGAAACGATTAACGTTCCGAAACGAGAGA 2678
 QY 559 CGATGATCTCAAGCCAGATGCTTGGTTTCTTGAAGAAACAGATGATGAGA---CCC 615
 Db 2679 AGCGTATCGCGATCAAAATCGGAAGTCTCAAAACAGATGAACCAAGGCTTC 2738
 QY 616 TGTGTTGTTGATGATGCTTAAACGCGGCTGCTTAACAGATGATGCTGCTATTC 675
 Db 2739 TGTGTTGATGATGATGATCAAAACGCGATCGCTTAATACATGATGCTGCTATTC 2798
 QY 676 GGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 726
 Db 2799 GGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2858
 QY 727 GGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
 Db 2859 GGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2918
 QY 787 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 846
 Db 2919 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2978
 QY 847 TTTGATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 906
 Db 2979 TTTGATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3038
 QY 907 GTGAGCTTTGCTGCTCCCTCCCTGATGATGATGATGATGATGATGATGATGATG 966
 Db 3039 GTGAGCTTTGCTGCTCCCTCCCTGATGATGATGATGATGATGATGATGATGATG 3098
 QY 967 GTTCTTTG----- 974
 Db 3099 GTTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3158
 QY 975 ----- 974
 Db 3159 AAAAATTCATTTTATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 3218
 QY 975 ----- 974
 Db 3219 AATTAATATGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3278
 QY 975 GGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1034
 Db 3279 GGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3338
 QY 1035 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1094
 Db 3339 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3398
 QY 1095 GAATCAAGAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1154
 Db 3399 AGCATCAAGAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3458
 QY 1155 GAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1214
 Db 3459 GAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3518
 QY 1215 AACAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1274
 Db 3519 CAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3578
 QY 1275 GTGAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1334
 Db 3579 GTGAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3638
 QY 1335 GTTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1394
 Db 3639 TTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3698

QY 1395 ACCCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1454
 Db 3699 ACCCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3758
 QY 1455 GTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1511
 Db 3759 GTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3818
 QY 1512 CGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1571
 Db 3819 CGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3878
 QY 1572 ATTGCGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1591
 Db 3879 ATTGCGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3898

RESULT 6
 US-09-349-385-7
 ; Sequence 7, Application us/09349385
 ; Patent No. US20020152495A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ico, Toshio
 ; APPLICANT: Fromm, Michael
 ; APPLICANT: Meyerowitz, Elliot
 ; TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT
 ; FILE REFERENCE: MBI-0002
 ; CURRENT APPLICATION NUMBER: US/09/349,385
 ; EARLIER FILING DATE: 1999-07-09
 ; EARLIER APPLICATION NUMBER: 60/115,967
 ; EARLIER FILING DATE: 1999-01-15
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 7
 ; LENGTH: 1611
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: cDNA
 US-09-349-385-7

Query Match 40.9%; Score 651.6; DB 10; Length 1611;
 Best Local Similarity 66.5%; Pred. No. 1e-210;
 Matches 1025; Conservative 0; Mismatches 484; Indels 33; Gaps 5;

QY 61 AGCGAAGCAACCTTGGCTTCTCTCTCTGCGCGTCAACATCATGCTGCGCATATCT 120
 Db 74 ACCGAAAGAAACCTTGGCAAGGCTCTTTGATCATAGGCTCATCTGCTTAACTACT 133
 QY 121 CTCTCTTATGACCTATCCCGGTGACTGCTTGGGAAATACCTTTGGCCGGTTA 180
 Db 134 TCTATTAATGATCAACCCGCTGCTCTGCTGCGGCAAGTACTACAC-----TA 186
 QY 181 ATATCCGCTTATTAAGAAACCGGAAGCTTATCCCGTCCAAAGGCTCCCTTGGT 240
 Db 187 CTCTCCCCCTTTCA-----TCAATCCCGTCCAAAGGCTTCTCTTATT 235
 QY 241 GGAAGATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
 Db 236 GGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 292
 QY 301 AAATTCGAGGCAAGGCTCATGCTTCAAGCTTGAAGAGATGCGGATGATGATGATGATG 360
 Db 293 AATTCGAGGCAAGGCTCATGCTTCAAGCTTGAAGAGATGCGGATGATGATGATGATGATG 352
 QY 361 TGCATCCGAGGATGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 420
 Db 353 TGCATCCGAGGATGAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 412
 QY 421 AAAGATGCGGCTTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
 Db 413 AAAGATGCGGCTTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 472

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.inpb

Page 8

QY 643 GCCTGCTTACACATGATGCTCTGTTATTCGACAGAGATGAGCTTGAACAAAAC 702
Db 730 GCGTCACTGAAACATGATGCTGCTGTTGACAAAGATACACCTTGACAGATA 789
QY 703 CA-----TGTGAGTTACGTGAATGCTGAAAGAGTTATGCTTGGCGAAG 753
Db 790 AACACGCAATGACGAGCTATCCATGTTGGTGAACAAAGCTATGACTGTGGGAC 849
QY 754 TTGATTTGACATGATCACTTCTTGGCTATCGAGTTGATCTTCAAAAGCTCGGCT 813
Db 850 CTTAATTTGGGAGAACATATCTTCTTGAAGAGTTGACCTACAGAAATCCGATTC 909
QY 814 AGATGTTCCACTGCTGACCAAGATTAACCGTTTGTATCCGATTAATCCGACAC 873
Db 910 ACCGTCTCCAAATTAATGCTCTCAAGTGAACCGTTGTTGTTCAATCATGCGACAC 969
QY 874 CGTATGAAACCGGTGATTTGCTGATTTGCTGAGCTTTGCTCTCCCTCATGCT 933
Db 970 CAGGCGACACAAACCAACCAACCGGATTTGCTGATGTTGCTCTCTCCAAAGT 1029
QY 934 TCAATTAATTTATCCAGCCGACATATGCGGCTTCTTGGAGATGATATTCAGAGA 993
Db 1030 CCGATTAATTTGCTCACTCTGACATGATTTGCTGCTGCGAAATGATTTAGGGG 1089
QY 994 ACAGACACAGTTGCTGTTAATGAGTGAATCTGCTGATGATGCTTCTTATCCAGAT 1053
Db 1090 ACCGACACGCTGCTGTTGATGATGATGATGATGATGATGATGATGATGATG 1149
QY 1054 ATGCAATCAACGCTGACAAACGAGCTGATGATGATGATGATGATGATGATGAT 1113
Db 1150 GTGCAAGGAGATGACAGAGATGATGATGATGATGATGATGATGATGATGATG 1206
QY 1114 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1173
Db 1207 GAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1286
QY 1174 CATCTTCAAGCCGCTTATCATGAGCCGCTTGGCCATTAACAGACATGATGAT 1233
Db 1267 CACCGCGCGGCGCTTCTCTGAGGCGCGCTTGGCCATTAACAGACATGATGAT 1326
QY 1234 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1293
Db 1327 GGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1386
QY 1294 CCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1353
Db 1387 CCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1443
QY 1354 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1413
Db 1444 ---AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1414 CGATTTGCGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGAT 1473
Db 1501 AGAACAATGCGCGGAGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1474 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1533
Db 1561 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
QY 1534 AGGCTTTCTTGCAGATGATGATGATGATGATGATGATGATGATGATGAT 1593
Db 1621 AGGCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680

RESULT 8
US-10-163-198-24
Sequence 24, Application US/10163198
Publication No. US20030126645A1
GENERAL INFORMATION:
Applicant: Rebecca E. Cahoon
Applicant: Elmer P. Heppard

APPLICANT: No. US20030126645A1uhio Nagasawa
APPLICANT: Hajime Sakai
TITLE OF INVENTION: Alteration Of Embryo/Endosperm Size During Seed Development
FILE REFERENCE: B01487 US NA
CURRENT APPLICATION NUMBER: US/10/163,198
PRIOR FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 60/295,921
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 60/334,317
PRIOR FILING DATE: 2001-11-28
NUMBER OF SEQ ID NOS: 111
SOFTWARE: Microsoft Office 97
SEQ ID NO 24
LENGTH: 1905
TYPE: DNA
ORGANISM: Glycine max
US-10-163-198-24
Query Match 36.8%; Score 585.8; DB 15; Length 1905;
Best Local Similarity 66.1%; Pred. No. 3.4e-188;
Matches 934; Conservative 0; Mismatches 452; Indels 27; Gaps 5;
QY 193 TCAAAACCGGAAAGTATATCCGCTGCAAAAGCTTCCCTTGTGAGAGATGCA 252
Db 294 TCCAAACACCTTAATATATCCGCTGCAAAAGCTTCCCTTGTGAGAGATGCA 353
QY 253 CTGATGCAAGCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 312
Db 354 CTGATGCAAGCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 410
QY 313 AAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 372
Db 411 ACCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 470
QY 373 GTAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 432
Db 471 GTAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 530
QY 433 TACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
Db 531 TACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 590
QY 493 CTTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 552
Db 591 CTTCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 650
QY 553 CAACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 612
Db 651 CAGCGCGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 704
QY 613 CCGCTGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 672
Db 705 GGTTCGAAATCCGAGGCTTCAAGAGACGCTGCTCAACACATGATGATGAT 764
QY 673 TTGCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 733
Db 765 TTGCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 824
QY 724 GTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 783
Db 825 GTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 884
QY 784 TCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 843
Db 885 AAGACTTGAATTTCAAAATATCGGTTTACGCTGCTCAAACTGCTCCCAAGTAA 944
QY 844 CCGTTGTTATCCGATTAATATCCGATTAATATCCGATTAATATCCGATTAAT 903
Db 945 CCGTTGTTATCCGATTAATATCCGATTAATATCCGATTAATATCCGATTAAT 1004
QY 904 TTGCTGACAGTTTGTCTCTCCCTCATGATGATGATGATGATGATGATGATGAT 963
Db 1005 TTGCTGACAGTTTGTCTCTCCCTCATGATGATGATGATGATGATGATGATGAT 1064

QY 964 GCCGTTCTTTGGAGATGATATTCAGAGAAACAGACAGTTCGGCTCTTAATCCAGTGG 1023
Db 1065 GCTGTCCTCTGGAGAAATGATATTTAGGGGACCGACCGATGGGTTTGTATGATGGG 1124
QY 1024 ATCCGCTAGATAGTGTCTTCAATCCAGATATGCAATCAACGCTACAAAACGAGTGTAT 1083
Db 1125 ATATGCGAAGATAGTGTCTTCAATCCAGAGTACAAAGAGGGTGCAGAGAGGCTGGAC 1184
QY 1084 CAAGTACT---CGGAATCAAGAGCCCTAGATGAATCTGACTTGGCTTCACTTCAAT 1140
Db 1185 GCGGTGTTGAGGTGTGTGGCGCGCTTTGAAAGAGAGACGTTGCGGAGCGGGTAT 1244
QY 1141 CTACGAGCTGTGGTGAAGAAAGATATGAGGCTTCACTCCAGAGCCCACTTATCATG 1200
Db 1245 CTTCGCGGTGTGGTGAAGAGGTTCTGAGGCTGCACTCCAGGCGGCTTCTGTGTG 1304
QY 1201 GCCCGTTTGGCATTAACAGACAGATGTTGATGTCTTGTTCGCGAGAGACCA 1260
Db 1305 GCCCGTTTGGCATTAACAGACAGATGTTGATGTCTTGTTCGCGAGAGACCA 1364
QY 1261 GCATGTGTAACATGTGGGCGGTATGCGATGATCAACGTTGGGTGATCTTTGGAG 1320
Db 1365 GCCATGTGTATATGTGGGCGGTATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1424
QY 1321 TTTAAACCTGAGAGGTTCTGTGCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1425 TTTAAACCTGAGAGGTTCTGTGCAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1478
QY 1381 GATTTGAGACTTGACCTTTTCGGGTGGGTCTGTGATTTGCCCGGAGAGAGAGAGAG 1440
Db 1479 GATTTGAGACTTGACCTTTTCGGGTGGGTCTGTGATTTGCCCGGAGAGAGAGAGAG 1538
QY 1441 TTTACTACCTTATGTTTGGACCGCGATGATGTTTACATGATTTGAATGGAGACCGTCC 1500
Db 1539 TTTAGCACCTGATCTTTTCGGGTGGGTCTGTGATTTGAATGGAGACCGTCC 1598
QY 1501 GATGTGACGCGCTTGACTTATCTGAGAAAGTGAAGGCTTTTCGCGAGATGCTTAATCC 1560
Db 1599 GATGTGACGCGCTTGACTTATCTGAGAAAGTGAAGGCTTTTCGCGAGATGCTTAATCC 1658
QY 1561 CTTCCTGCTAATGCGCGGCTGAGCGGAGTTAA 1593
Db 1659 CTCTATGTTAAAGTTCCGCCCTAGCGGTGATTA 1691

RESULT 9

US-09-349-385-13

Sequence 13, Application US/09349385

Patent No. US20020152495A1

GENERAL INFORMATION:

APPLICANT: Ito, Toshiro

APPLICANT: Fromm, Michael

APPLICANT: Meyerowitz, Elliot

TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT

FILE REFERENCE: MBI-0002

CURRENT APPLICATION NUMBER: US/09/349,385

EARLIER FILING DATE: 1999-07-09

EARLIER APPLICATION NUMBER: 60/115,967

NUMBER OF SEQ ID NOS: 13

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 13

LENGTH: 1799

TYPE: DNA

ORGANISM: Phalaenopsis sp.

FEATURE:

OTHER INFORMATION: cDNA

FEATURE:

OTHER INFORMATION: Hybrid SM9108

US-09-349-385-13

Query Match

22.3%; Score 355; DB 10; Length 1799;

Best Local Similarity 58.0%; Pred. No. 1,9e-109; Matches 744; Conservative 0; Mismatches 500; Indels 39; Gaps 5;

QY 304 TTGGAGCGAAGAGGCTTCATGCGCTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAG 363
Db 273 TTGGAGCGCTCCGCTTCATGCGCTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAG 332
QY 364 AATCCGAGCTGAG 423
Db 333 CAACGAG 392
QY 424 GATGAGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 483
Db 393 GATGAGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 452
QY 484 TGGAGAGCGCTTCCGCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATG 543
Db 453 TGGAGAGCGCTTCCGCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATG 512
QY 544 GCCGAGCGCAACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 601
Db 513 TTTGAGAGAGCAACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 572
QY 602 -GTAGTAAGAACCTGTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 660
Db 573 GAG 632
QY 661 ATGTCCTGTATTTGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 633 ATGTCCTGTATTTGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 692
QY 721 ATGTCCTGTATTTGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 777
Db 693 ATGTCCTGTATTTGGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 752
QY 778 TGGCTATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 837
Db 753 TGGCTATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 812
QY 838 GTAAACCGTTTGTATCCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 885
Db 813 GTAAACCGTTTGTATCCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 872
QY 886 -----GGTATTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 939
Db 873 ATGATGAG 932
QY 940 AATATTCGACCGGACATATATGCGCTTCTTTGGAGAGAGAGAGAGAGAGAGAGAGAG 999
Db 933 AGACTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 992
QY 1000 ACAGTTGCGGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1059
Db 993 ACAGTTGCGGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1052
QY 1060 TCAACGCTAACAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1119
Db 1053 TCAACGCTAACAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1112
QY 1120 GATTTGCTTCACTTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 1179
Db 1113 GATTTGCTTCACTTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 1172
QY 1180 CCAAGCCCATCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1239
Db 1173 CCAAGCCCATCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1232
QY 1240 CTGTTCCGCGAG 1299
Db 1233 ATGATTCCTGCTGAG 1292
QY 1300 GTGTGGGTTGATCTTTGAGATTTAACTGAGAGGTTGTTGGCAAAAGAGTGAAGTGG 1359

US-09-349-385-9
 ; Sequence 9, Application US/09349385
 ; Patent No. US20020152495A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ito, Toshio
 ; APPLICANT: Fromm, Michael
 ; APPLICANT: Meyerowitz, Elliot
 ; TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT
 ; FILE REFERENCE: MBI-0002
 ; CURRENT APPLICATION NUMBER: US/09/349,385
 ; EARLIER FILING DATE: 1999-07-09
 ; EARLIER APPLICATION NUMBER: 60/115,967
 ; EARLIER FILING DATE: 1999-01-15
 ; NUMBER OF SEQ. ID NOS: 13
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 1990
 ; TYPE: DNA
 ; ORGANISM: Pinus radiata
 ; FEATURE:
 ; OTHER INFORMATION: CDNA
 ; US-09-349-385-9

Query Match 21.3%; Score 348.2; DB 10; Length 1990;
 Best Local Similarity 55.5%; Pred. No. 4.3e-107;
 Matches 768; Conservative 0; Mismatches 588; Indels 27; Gaps 4;

QY 211 ATTCCGGTCCAAAGGCTCCCTTGGTGAAGCATGTCATGTCAGCACTCTA 270
 DB 307 ATTCTGCGCCCAAGGATGCGGATTAATGAAAGTCTAATGACATGAGCGCTT 366
 QY 271 GCTCACCAGCAATCGCTGATGCAAGTGAAGAAATTCGAGCCAGAGGCTCATGCTTTC 330
 DB 367 CTCACCGCAAGCTGATCTCTGCTCGCTTCATGAGTGCAGAGCAAGCTCATGCTTTC 426
 QY 331 AGCTTAGAGAGACTCGCGTATGTCACGTCGATCCGAGTACCGAAGAGATTCG 390
 DB 427 AGCTGGGCTGACGCGCGCTGTTATATACGTCGATCCGAGTGGCCGAGGTTGCTC 486
 QY 391 AATAGCCGGTCTTGGTGAATGCAAGCGTTAAAGAAATCGCTTACTCATGATGTTTAA 450
 DB 487 ACCGCCCTCACTTCGCAAGACAGCCCTCAAGCAATCCGACAGAGTGTGCTTGG 546
 QY 451 AGAGCAATGGTTTTCACACACAGGTTTACTGCGAAGCGTTGCCGTTATCGCTTGG 510
 DB 547 AGGCTATGCTGCTGCAAGCCGAGGCTGCTGAGGCTGCTCAGAGAAATGCTTCT 606
 QY 511 AACCATCTCTTATGATCAAAACAAATCAGAAAGCCGAGACGACGAGTATCTCA 570
 DB 607 GCTCATCTCTTGTCTCTCCAGAGATGCTGCCCATGAAAGCTGGCCGACGCGCAAGTT 666
 QY 571 AGCCAGATGTTGATGTTTCTTGAAGAAACAGATGATACAGACCTGTTT---GTTCTG 627
 DB 667 GTGCAATGCTGATGATGATCAAAAGAGATCATCTTAAGGCTGTTGTAAGGTAAG 726
 QY 628 GAGTGTCTTAAACGAGGCTGCTTAAACAATGATGCTGCTGATTCGAGCAAGATAT 687
 DB 727 CGACACTGCAAGAGAGGCGCTGATACATATCTGGAAGAGCGCTTGGAAAGAGTTTC 786
 QY 688 GAGCTGTAAGAAACCATGTTG-----AGTTACGTAAATGTCGAAAGAGTTAT 738
 DB 787 GACATGCTCTCATGAAATGAAAGATGAAAGAGCTGCGGAGATGTTGTAAGGTTTC 846
 QY 739 GATTGCTCGGAAGTGAATGGAATGATCATCTTCTTGGCTATCGAGTTGATTCCT 788
 DB 847 CAATCTGTAAGAGCTTCACTGAGCGGATATCTCCATGAGCTGCGACCGCTGACCT 906
 QY 799 CAAGAGCTCGGCTAGATGTTTCCACTGCTACCAAGGTAAACCGGTTGATTCGG 858
 DB 907 CTTCGAATCCATGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
 QY 859 ATTATATCCGAACCGTA-----ATCAACCGTGAATTTGCTCGGATTC 906

DB 967 ATCATCGAACGACCGCTCGGAAAGAGCAGCTGAGAAAGCGCGCATGATGATTT 1026
 QY 907 GTCAAGCTTTTGTCTCTCCCTCCATGTTCAATTAATACCGACCGCGCATATATGCC 966
 DB 1027 GTTACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1086
 QY 967 GTTCTTTGGAGATGATTAATCAGAGAAACAGACAGCTTGGCTTATATGATGATC 1026
 DB 1087 GTTCTTGGAGATGATTAATCAGAGAAACAGACAGCTTGGCTTATATGATGATC 1146
 QY 1027 CTGCTAGATGATGATTAATCAGAGAAACAGACAGCTTGGCTTATATGATGATC 1086
 DB 1147 ATGACAGAACTGCTTTCACAGCTTGAAGCTGAGAAAGCTGAGAGATGAGAGCC 1206
 QY 1087 GTAGCGGAAATCAAGAGCCCTGATGAAATCTGATCTTGGCTTATATGATGATC 1146
 DB 1207 GTAGTGGGGACAGACCGAAGCTGAGATTCAGACATTCGAAAGCTGCGTATCA 1266
 QY 1147 GCTGTGTAAGAAAGATTAATGAGCTTATCTTCCAGCCCACTTCTATATGAGCCGCT 1206
 DB 1267 GAGTGTGTAAGAAAGAGCTGCTGAGATGATCACTGAGACCGCTCTATATATGAGCTGCT 1326
 QY 1207 TTGGCAATTAAGA---CAGATGCTGATGATGCTGCTTGTCCGAGAGGACACAGCA 1263
 DB 1327 CTCTTACAGAAAGATGATCAATGAGCGGATGAGATGATGATGATGATGATGATGATGAT 1386
 QY 1264 ATGTGAACATGTGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1323
 DB 1387 ATGTGAATATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
 QY 1324 AAACCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1383
 DB 1447 GCTCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1506
 QY 1384 TTGAGACTTGAACCTTTCGAGTCCGCTGCTGATGATGATGATGATGATGATGATGAT 1443
 DB 1507 CTGAGGCTTGGCTTTCGAGTCCGCTGCTGATGATGATGATGATGATGATGATGAT 1566
 QY 1444 ACTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1503
 DB 1567 GCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1626
 QY 1504 GATAAGCGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1563
 DB 1627 GAACTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1686
 QY 1564 CCT 1566
 DB 1687 CAT 1689

RESULT 12
 US-10-163-198-22
 ; Sequence 22, Application US/10163198
 ; Publication No. US20030126645A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rebecca E. Cahoon
 ; APPLICANT: Elmer P. Heppard
 ; APPLICANT: No. US20030126645A1
 ; APPLICANT: Hajime Sakai
 ; TITLE OF INVENTION: Alteration of Embryo/Endosperm Size During Seed Development
 ; FILE REFERENCE: BB1487 US NA
 ; CURRENT APPLICATION NUMBER: US/10/163,198
 ; EARLIER FILING DATE: 2002-06-05
 ; EARLIER APPLICATION NUMBER: 60/295,921
 ; EARLIER FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: 60/334,317
 ; NUMBER OF SEQ. ID NOS: 111
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 22
 ; LENGTH: 1934
 ; TYPE: DNA

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rmpb

Page 12

ORGANISM: Glycine max
US-10-163-198-22

Query Match 19.9%; Score 316.6; DB 15; Length 1934;
Best Local Similarity 55.1%; Pred. No. 2.6e-96;
Matches 750; Conservative 0; Mismatches 569; Indels 42; Gaps 5;

254 TCATGTCAAGCACTCTAGCTCACCAGCAATCGCTGATGCACTGAGAAATTCGAGCCA 313
255 TCTTACTAGGCTCCACGCGCTCACCCTGCTTATCAACATCGCCCTGATACCAACCGG 314
314 AGAGGCTATGAGCTTTGAGCTTAGAGAGACTCGCGATGCTGACGTCGATCCGAGC 373
315 AAAAAGTATGAGCTTTGAGCTTAGAGAGACTCGCGATGCTGACGTCGATCCGAGC 374
374 TAGCGAAAGATTCGATAGTCCCGGTTTTCGCTGATGCAACCGGTTAAAGATCGCTT 433
375 CCGCTAAGAGATTCGAGCTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 434
434 ACTCATGATGTTTAAAGAGAGATTCGAGCTAGGAGAGAGAGAGAGAGAGAGAG 435
435 ATGAGCTTCTCTCCACCGCGCAATGAGGTTTGAACGCTATGAGGAGATTCGAGAG 494
494 TTGCGCGATGCTTCAACCATCTCTTATGATCAAAACAATGAGAGAGAGAGAGAG 553
495 TGAGAGAGATTCGAG 554
554 AAGCAGAGATTCGAG 610
555 TCAG 614
611 AACCGTTTTCGAG 670
615 AACAGTTGAGTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674
671 TATTCGAG 727
675 TCTTTGATGAGTTTATGAGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 734
728 AAGAGTTATGATTTGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 787
735 GTGAG 794
788 AGTTGATCTCAAG 847
795 GGTGATTTGAG 854
848 TTGTATCCCGAGTTATATCCGAAACCGGATCAACCGG----- 887
888 -TGATTTGCTCGATTTCTGAGAGTTTCTCTCCCTCATGCTTCAATTAATAT 946
915 ATGAG 974
947 CCGAGCGGAGATATCCGCTTTTGGAGAGAGATTCAGAGAGAGAGAGAGAGAGAG 1006
975 GTGAG 1034
1007 CCGTCTTAATCGAGAGATTCCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1066
1035 CAATTCGTAG 1094
1067 TACAAAG 1126
1095 CAGAGCGGAG 1154
1127 CTTCATCTCATATCAACGAGCTGAGTGAAGAGAGATTCAGAGAGAGAGAGAGAG 1186
1155 CGAAGCGGAG 1214
1187 CACTTCTATGAG 1243

1215 CGCTACTCTCGTGGGCTCGCTTCGTGTCAGAGAGATTCAGAGAGAGAGAGAGAG 1274
1244 TTCCGAG 1303
1275 TTCCGAG 1334
1304 GGGTATGCTTTGAG 1363
1335 GGGCGAG 1382
1364 TTTCGAG 1423
1383 TGAGATATGAG 1442
1424 CCGGAG 1483
1443 CTGGAG 1502
1484 TTGAATGAG 1543
1503 TTGATTTGAG 1562
1544 GCGAGATGAG 1584
1563 TGAGATGAG 1603

RESULT 13
US-10-163-198-40
Sequence 40, Application US/10163198
Publication No. US20030126645A1
GENERAL INFORMATION:
APPLICANT: Rebecca E. Cahoon
APPLICANT: Bimer P. Heppard
APPLICANT: No. US20030126645A1uhiro Nagasawa
APPLICANT: Hajime Sakai
TITLE OF INVENTION: Alteration of Embryo/Endosperm Size During Seed Development
FILE REFERENCE: B1487 US NA
CURRENT APPLICATION NUMBER: US/10/163, 198
CURRENT FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 60/295, 921
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 60/334, 317
PRIOR FILING DATE: 2001-11-28
NUMBER OF SEQ ID NOS: 111
SOFTWARE: Microsoft Office 97
SEQ ID NO 40
LENGTH: 1687
TYPE: DNA
ORGANISM: Alstroemeria caryophylla
US-10-163-198-40
Query Match 18.2%; Score 290.6; DB 15; Length 1687;
Best Local Similarity 54.1%; Pred. No. 1.7e-87;
Matches 712; Conservative 0; Mismatches 579; Indels 24; Gaps 5;

```

Db 433 GCGGCACTCCGCGCACCCACCTCTTCAGCCGCCAGGCGCTCCGCCCTCCGCCCTCC 492
Qy 557 GAGCACTGATCTCAAGCCAGATGATGATTTCTTGAAGAAAGAGATAGTAAAGACCT 616
Db 493 GCGGCACTCCGCGCACCCACCTCTTCAGCCGCCAGGCGCTCCGCCCTCCGCCCTCC 552
Qy 617 GTTTGCTGCTGAG---TTGCTTAAACCGCGCTGCTTAAACAGATGATGCTCTGAT 673
Db 553 AGGTGAGATCAAGCGCGCTCTCCACTGCGCTCCCTCAACAGATGAGCGAGCGTGT 612
Qy 674 TCGGACAAGATGATGAGCTTGAAGAAA---ACCATGATGATGATGATGATGATGAT 730
Db 613 TCGGCGCGCTCAAGCTTCCGACCCCGGAGCGCGCTCGAGCTCGAGCGCTTGTGAG 672
Qy 731 AAGGTTATGATTTGCTCGGAAAGCTGTAATGATGATGATGATGATGATGATGAT 790
Db 673 AAGGTTATGATTTGCTCGGAAAGCTGTAATGATGATGATGATGATGATGATGAT 732
Qy 791 TTGATCTCTAAAGCTCCGCTTATGATGATGATGATGATGATGATGATGATGATGAT 850
Db 733 TTGATCTCTAAAGCTCCGCTTATGATGATGATGATGATGATGATGATGATGATGAT 792
Qy 851 TATCCGCGATTTATTCGGAACAC---GTATCAAAACCGGTGATTTGCTCGATTTG 907
Db 793 TCGGCGCGCTCAAGCTTCCGACCCCGGAGCGCGCTCGAGCTCGAGCGCTTGTGAG 852
Qy 908 TCGAGCTTTGCTCTCCCTCCAGTGTTCAGATTAATTTCCGACCCGAGATTAATGCG 967
Db 853 CCGTAGACTTCGTGAGCTGCTGCTCCAGCAAGAGCTTCGATTCGATGATGATGATGAT 912
Qy 968 TTTCTTTGGGAGATATTTTCAGAGAAAGAGACAGTTCGAGTTCGATGATGATGATGAT 1027
Db 913 TTTCTTTGGGAGATATTTTCAGAGAAAGAGACAGTTCGAGTTCGATGATGATGATGAT 972
Qy 1028 TCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
Db 973 TCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1032
Qy 1088 TATGTCGGAATTCAGAGAGCTTATGATGATGATGATGATGATGATGATGATGATGAT 1147
Db 1033 TTTGCGGCGGTGAGAGGCTGCTGCGCGAGCTGCGAGCTGCGAGCTTCTTACTTCAAG 1092
Qy 1148 CTGTGTGAAAGAGATTTAGGCTTCACTCTCGAGGCCACTTCTTATCATGAGCGCTT 1207
Db 1093 CCAATGTCAGAGAGTGTGCTGAGATGAGACCCCGCGCGCTGCTCTCTGAGCTGCGC 1152
Qy 1208 TGGCCATTAACAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1267
Db 1153 TCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
Qy 1268 TGAACATGTCGCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1327
Db 1213 TGAACATGTCGCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
Qy 1328 CTGAGAGGTTGTCGCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1387
Db 1273 CGGAGAGGTTGTCGCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1320
Qy 1388 GACTTGAACCTTTGCGGTCGCGCTGCTGCGATTTGCGCGGAGAGATTTGTTTACTA 1447
Db 1321 GCGTGGCGCGCTTCCGCTGCGGAGAGAGTGTGTCCGCGCAAGCGATGAGGCTGCGCA 1380
Qy 1448 CCGTTATGTTTGAAGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1507
Db 1381 CCGCGCATCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1437
Qy 1508 ACGGCGTGAATTTATGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1562
Db 1438 ATGCGGTTATCTCTCGGAAAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1492

```

RESULT 14

```

US-10-163-198-96
; Sequence 96, Application US/10163198
; Publication No. US2003012645A1
; GENERAL INFORMATION:
; APPLICANT: Rebecca E. Cahoon
; APPLICANT: Elmer P. Heppard
; APPLICANT: No. US2003012645A1unhiro Nagasawa
; TITLE OF INVENTION: Alteration Of Embryo/Endosperm Size During Seed Development
; FILE REFERENCE: B01487 US NA
; CURRENT FILING DATE: 2002-06-05
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: 60/1295,921
; PRIOR FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 96
; LENGTH: 1545
; TYPE: DNA
; ORGANISM: Zea mays
US-10-163-198-96

Query Match      18.0%; Score 286; DB 15; Length 1545;
Best Local Similarity 54.1%; Pred. No. 6,1e-86;
Matches 692; Conservative 0; Mismatches 550; Indels 36; Gaps 4;

Qy 306 CCGAGCCAGAGAGCTCATGAGCTTTCAGCTTGAAGAGTACGCTGATGATGATGATGAT 365
Db 255 CCGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 314
Qy 366 TCCGACGATGAGGAGAGATTTGATAGCCCGGTTTTCGATGATGATGATGATGATGAT 425
Db 315 GCGGAGACCGGCGCGGAGCTCTGCGAGCGCGCTTCCGAGCGCGCGCGCGCGCGCGAG 374
Qy 426 ATGCGCTTACTACGATGATTTTAAAGAGCAATGCTTTCAGACAGAGGATTTACTG 485
Db 375 CCGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 434
Qy 486 GCGAGCTTCCGCTATGCTTTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 545
Db 435 GCGGCGCTTTCGCGCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 494
Qy 546 CAGAGCGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 605
Db 495 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 554
Qy 606 TAAAGAACCTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 665
Db 555 ACGAGAGGTC---GTATGCGCGCGCGCTGCTGCGAGCGCGCGCGCGCGCGCGCGCGCG 611
Qy 666 CTCTGATTTTCGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722
Db 612 CACCGGTTCGCGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
Qy 723 GTCGAGAGAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 782
Db 672 GGTGAGAGAGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 731
Qy 783 ATGAGATTTGATCTTCAAGAGCTCGGCTGATGATGATGATGATGATGATGATGATGAT 842
Db 732 CAGTGGCTGAGCTGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAG 791
Qy 843 CCGGTTGATATCCGAGTTATATCCGAGACCGGTAATCAACCGGATTTGCTGCTGTA 902
Db 792 CGTGTTCGCGCGAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 851
Qy 903 T-----TTGTCGAGCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 944
Db 852 GCGGCGCGCGCGCGAGAGCTTCTGAGAGCTTCTGCTGAGAGCTGAGAGGCGAGAGAGCT 911
Qy 945 ATCCGACCGGAGATATAGTGTCTTTGGAGATGATGATTTAGAGAGAGAGAGAGAGCT 1004

```

Db 912 GTGGATCTCGACATGATTCGCTGCTCTTGGAGATATCTTTCAGAGGACCAACGCT 971
 Qy 1005 TGGCGTCTTAATCGAGTGAATCCGCTAGAGATGCTTCATTCAGATATGCAATCAAC 1064
 Db 972 GGCGATCTGTGAGTGGGATGAGCGCGAGATGCTGTCACCCGGCATTCAGTCCAA 1031
 Qy 1065 GGTACAAAACGAGTGTGATCAAGTATCGGAAATCAAGACCCCTAGATGAATCTGACT 1124
 Db 1032 GGCACAGCGAGAGCTGACCGCTGTGTGGCGCGCGCCGCTTTGCGACGCGACGT 1091
 Qy 1125 GGCCTCACTTCATATCTAAACGCTGTGTGAAGAAATATAGCTTCATCTCCAGG 1184
 Db 1092 GCGCCGCTGTGCTTACCTGACGCGCTGTGTGAAGAGCGCTCCGCTGACCCCGCGG 1151
 Qy 1185 CCCACTTTCATATGAGCGCTTTGGCTTAACAGACACGATCGTTGATGTGTCTGT 1244
 Db 1152 CCGCTGCTCTGTGGGCGCGCTGTGCGGCGGCTGTGTGTGTGTGTGTGTGTGT 1211
 Qy 1245 TCCGACAGGACCAACGATGTGAACATGTGGCGGTATGCGATGATCCACAGTGTG 1304
 Db 1212 CCGCGCGGACACAGCGCATGTGTCAATGTGGCCATGCGCACGACCCCGCGGTGTG 1271
 Qy 1305 GATTGATCTTTGAGTTTAACTGAGAGTTGTGTGCAAAAAGAGTGAAGTGTGAGTT 1364
 Db 1272 GCGGAGCGCTTCGCTTTCGCGCGCGAGCGGTTCGAG-----GTGGAGAGCT 1319
 Qy 1365 TTGGTCTTGGGTGAGATTGAGACTTGCACCTTCGAGTGGGTGCTCGAATTTGCC 1424
 Db 1320 GAGCTGTGCGCGCGGACCTCCGCTTCGCGCTTCGCGCGCGCGCGGTGTGCC 1379
 Qy 1425 CGGAGAGATCTTGTGTTTACTACCTGATATGTTTGAACGCGATGATGATGATG 1484
 Db 1380 GGGCAAGACGTGGCGCTCGCCATGTCACCTTCGCTCGCGACGCTGTCACCGCT 1439
 Qy 1485 TGAATGGGAGCCGTCGATGATGAAGGAGTTGACTATCTGAGAACTGAGGCTTTCTG 1544
 Db 1440 CCGGTGGCGCGCGCGAGCGCGCGCGGTGCACTTGGCGAGCGCTTCGCGATCTCC 1499
 Qy 1545 CGAGATGAGTAATCTCT 1562
 Db 1500 GAGATGAGAGAGCCCT 1517

RESULT 15
 US-10-163-198-94
 ; Sequence 94, Application US/10163198
 ; Publication No. US20030126645A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rebecca E. Heppard
 ; APPLICANT: Elmer P. Heppard
 ; APPLICANT: No. US20030126645A1unhio Nagasawa
 ; APPLICANT: Hajime Sakai
 ; TITLE OF INVENTION: Alteration Of Embryo/Endosperm Size During Seed Development
 ; FILE REFERENCE: B1487 US NA
 ; CURRENT APPLICATION NUMBER: US/10/163,198
 ; CURRENT FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: 60/295,921
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: 60/334,317
 ; PRIOR FILING DATE: 2001-11-28
 ; NUMBER OF SEQ ID NOS: 111
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 94
 ; LENGTH: 1758
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; US-10-163-198-94
 Query Match 17.8%; Score 283; DB 15; Length 1758;
 Best Local Similarity 55.1%; Pred. No. 7e-85;
 Matches 709; Conservative 0; Mismatches 530; Indels 48; Gaps 6;

Qy 306 CGAGCCCAAGAGCTATGCTTTCAAGCTTAGAGAGACTCGCGTATTCGTACGTGCA 365
 Db 444 CGGCGCAAGGACATGATGGCTTCTCGGTGGGCTGACCGGTCTGTGTGTGACCA 503
 Qy 366 TCCGACGTACGAAAGATTCGAAATAGCGGGTTTGTGTGATGACCGGTTAAGA 425
 Db 504 GCGCATAGCGGCGGAGATCTCGACGCGCGCTTGTGGGCAACGCGCGCTCAAGA 563
 Qy 426 ATCGCTTACTACTGATGTTTAAAGAGCAATGTTTTCACCAACG--TGTTTA 482
 Db 564 CGGCGGCGCACACGTCTTTCACCGCGCATGGCTTCGCGCTTCGAGAGAGCGCA 623
 Qy 483 CTGGGAAAGCTTGGCGGTATCGTTGAAACATCTCTTGTATCAAAACAAATCAAG 542
 Db 624 CTGGCGGCGCTCGCGCGCTCGCGCACACCTGTTGGCCCGCGCGCGCGCGG 683
 Qy 543 AGCGAGAGCGCAACGAGATCTCAAGCCAGATGATGATTTCTTGAAGAAACAGAG 602
 Db 684 TCCCGCACACACCGCGCTCATGCGGAGGCGATGATGCGCACGTGCGCGTGCAT 743
 Qy 603 TAGTAAAGAACCTGTTTGTGTGAGT---TGTTAAACGCGCTGCTTAACAAT 659
 Db 744 GCGCGCGCACCGGAGAGTCTCTCAAGCGCGGTGTGTGATGTGATCTTCAACAGT 803
 Qy 660 GATGTCTGTATTGCGACAAGATATGAGCTTGAAGAAACCAATGTGAGTACGTGA 719
 Db 804 CATGCGACACGTGTTTGGCAAGCTAGCAATGGGAGCGGAGGCGCGCTTCGGA 863
 Qy 720 A---ATGTGAAAGAGTTATGATTTGTCGAAAGTTGAATTGACATGACCTTCC 776
 Db 864 CGAGATGTGGCGCGAGGCTACGACCTCTGGGACAGTTCAACTGGCTGACCACTG 923
 Qy 777 TTGGCTATGAGATTTGATCTCAAGACTCGGCTGATGATGTTTCACTGTACAA 836
 Db 924 ATTGCTCAAGATCTGACATCCCGAGGCGGTGCGCGCGGTGACAGCGCTGTGAAA 983
 Qy 837 GGTAAACGGTTTGTATCCCGATATATATCCGAACCGTAATCAACCG-----887
 Db 984 GTGTGAATGTTGTGTGGCAAGATCATCTTGGAGCAAGAGGCGCGCGCAATGAGG 1043
 Qy 888 -----TGAATTGCTCGGATTTGCGGAGTTTGTCTCCCTCATGATTC 935
 Db 1044 AGTGTGGGAGATGAGTGAATGATGATGATGATGATGATGATGATGATGATG 1103
 Qy 936 AGATTAATTATCCGACCCCGACATAATCGCGCTTCTTGGAGATATATTCAGAGAAC 995
 Db 1104 GAGAGAGCTGACAGATGCGGACATGATCGCTTCTTGGAGATGATGATGATGATG 1163
 Qy 996 AGACACAGTTGCGGTCTTAATCGAGTGAATCCCTGAGAGTGTCTTCATCAGATAT 1055
 Db 1164 CGACACCGTGGCATCTTGAATGAGTGAATGAGTGAATGAGTGAATGAGTGAAT 1223
 Qy 1056 GCATCAACGATCAAAACGAGCTGATCAAGTATGCGGAAATCAAGAGCTTCAATGA 1115
 Db 1224 CAGAGGAGGCGGAGGCGGAGGCTGAGCGGCTGAGGAGATGCGGCGCGCTGCGG 1283
 Qy 1116 ATCTGACTTGGCTTCACTTCATATCTGAAGGCTGTGTGAAGAGTATGAAGCTTCA 1175
 Db 1284 CGCGAGGTGCGGAGCTTACCTTCAATCAGTGAATGATGATGAGAGAGCTGCGATGCA 1343
 Qy 1176 TCTCCAGGCGCACTTTCATATGAGGCGGCTTGGCCATAACAGACAGATGTTATGG 1235
 Db 1344 CCGCGAGCGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1403
 Qy 1236 TGTCTTGTTCGCGAGGACACAGCAATGTGAAATATGAGGCGGTATGAGATGATCC 1295
 Db 1404 CCACCTGTTCGCGCGGACACAGCAATGTGAAATATGAGGCGGTATGAGATGATCC 1463
 Qy 1296 ACAGGTGTGAGTATCTTGTGAGTTTAACTGAGAGGTTGTGTGCAAAAGAGTGA 1355
 Db 1464 GCGATCTGGGCGAGCGGAGAGTTCGCGCGGAGCGGTTCCAG-----GA 1511
 Qy 1356 GGTGAGTTTTCGCTTCTTGGGTGAGTTTGAAGCTTCCGCGGTGCGGTGCTGCG 1415

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rnpb

Db 1512 GGAGGAGCTGAGCGTCTCTGGGAGCGACCTCCCGCTGGCCCCCTTGCGCGCGCGGCGCG 1571
QY 1416 GATTGGCCCGGGAAGAATCTTGTTTACTACCGTTATGTTTGAACGGCGATGATGTT 1475
Db 1572 CGCCTGGCCCGCAAGATCTGGCCCTGGCACCCACCACTCTGGGTGCGCCGCTTCT 1631
QY 1476 ACATGAGTTGATGGGAGCCGTCCGATGTTAOGGGTGACTTATCTGAGAACTGAG 1535
Db 1632 GCACAGTTGAGTGGG-----CGCCGCGGGGGGTGACCTGTGAGAGGCTGAG 1685
QY 1536 GCTTCTTGGGAGATGACTAATCCTCT 1562
Db 1686 CATGTGGCTGGAGATGGCAAGCGGCT 1712

Search completed: January 17, 2004, 06:00:01
Job time : 544 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 05:51:02 ; Search time 5880 Seconds
(without alignments)

11083.169 Million cell updates/sec

Title: US-10-022-025a-2

Perfect score: 1593

Sequence: 1 atgctacgaactcgaag.....tgcgcctagcgcagttaa 1593

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size : 30

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_or:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rdd:*
36: em_htg_mam:*
37: em_htg_vrc:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	976	61.3	55870	8 AC005819	AC005819 Arabidops
C 2	976	61.3	84825	8 AC006418	AC006418 Arabidops
C 3	232	14.6	740	8 ATH552847	ATH552847 Arabidops
C 4	141	8.9	504	8 ATH553770	ATH553770 Arabidops

ALIGNMENTS

RESULT 1
AC005819/c 55870 bp DNA linear PLN 11-MAR-2002
LOCUS Arabidopsis thaliana chromosome 2 clone T3A4 map C1C06C03, complete
DEFINITION sequence.
AC005819
VERSION AC005819.3 GI:20197448
KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 55870)
Rounsley,S.D., Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Mason,T.M.,
Shen,M., Rensing,C.M., Fraser,C.M., Somerville,C.R. and Venter,J.C.
Unpublished
2 (bases 1 to 55870)
Lin,X.
Direct Submission
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
3 (bases 1 to 55870)
Town,C.D. and Kaul,S.
Direct Submission
Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, cdtown@igf.org
On Apr 18, 2002 this sequence version replaced gi:6598489.
COMMENT
FEATURES
source
1..55870
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/culturVar="Columbia"
/db_xref="taxon:3702"
/chromosome="2"
/map="C1C06C03"
/clone="T3A4"
1..14747
/note="overlap with BAC clone F13A10
(AC006418:70079..84825)."
92..1276
/gene="At2g46630"
/note="synonym: T3A4.1"
<92..>1276
/gene="At2g46630"
92..1276
/gene="At2g46630"
/product="putative extensin"
/protein_id="AAC6930.1"
/db_xref="gi:3831447"
/translation="MSNIRPWFRLSSIRAPTAQGSDDPPQPPPTSRSLVVRPAK
QSPPRQPSPPROQDPPSPROQDPLPPROKAPTSPOBRSYHSPSRHMS
PPTPKATPPPPPPRSYTSPPSKYQALPPKPPSPSPSAHSRSTTSSEVYKTR
SPSEENHRKAPSPRVLSPTISPAIILSEBETTKNLTLEKTSOTETHHNNHN
HDYNNHYNQNSYNQNNQNNPKKHROPSSSDSENIMSTRVITIAENKCAVM

Query Match 61.3%; Score 976; DB 8; Length 55870;
 Best Local Similarity 100.0%; Pred.No. 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 976; Conservative 0;

QY 1 ATGGCTACGAACTCGAAGCTCTTATCTTTCGCTTTGTCGAATGACGCTTCTA 60
 DB 9980 ATGGCTACGAACTCGAAGCTCTTATCTTTCGCTTTGTCGAATGACGCTTCTA 9921
 QY 61 AGCCAAACCACTTTCCTTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 120
 DB 9920 AGCCAAACCACTTTCCTTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 9861
 QY 121 CTCTTCTTATGACACTATCCCGGTGACCTGCTTGGGGGAAATACCTCTTCCGCGGTTA 180
 DB 9860 CTCTTCTTATGACACTATCCCGGTGACCTGCTTGGGGGAAATACCTCTTCCGCGGTTA 9801
 QY 181 ATATCCGCTTCAATCAAAACCGGAAACGTTATTCCTCCGCTTCCAAAGGCTTCCCTTGGTT 240
 DB 9800 ATATCCGCTTCAATCAAAACCGGAAACGTTATTCCTCCGCTTCCAAAGGCTTCCCTTGGTT 9741
 QY 241 GGAAGCATCTCACTCATGTCAAGCACTCTAGCTCAACGCAATCGCTGATGAGCTGAG 300
 DB 9740 GGAAGCATCTCACTCATGTCAAGCACTCTAGCTCAACGCAATCGCTGATGAGCTGAG 9681
 QY 301 AAATTCGAGCCAAAGAGCTCATGCTTTTCAAGGAGACTGCTGCTGATGCTGATG 360
 DB 9680 AAATTCGAGCCAAAGAGCTCATGCTTTTCAAGGAGACTGCTGCTGATGCTGATG 9621
 QY 361 TCGAATCCGACGTAAGGAAAGATCTGTAATAGCCGCTTTTGTGATGACGCGGTT 420
 DB 9620 TCGAATCCGACGTAAGGAAAGATCTGTAATAGCCGCTTTTGTGATGACGCGGTT 9561
 QY 421 AAGAAATCGGCTTACTCATGATGTTTACAGAGCAATGTTTTCACCAACGCGTGT 480
 DB 9560 AAGAAATCGGCTTACTCATGATGTTTACAGAGCAATGTTTTCACCAACGCGTGT 9501
 QY 481 TACTGCGGAAACGCTTCCCGTATCGCTTGAACCATCTCTTATGTAACAAACGATCAG 540
 DB 9500 TACTGCGGAAACGCTTCCCGTATCGCTTGAACCATCTCTTATGTAACAAACGATCAG 9441
 QY 541 AGAGCCGAGACGCAACGAGATGATCTCAAGCCAGATGCTTGAAGTTCTTGAAGAAACAG 600
 DB 9440 AGAGCCGAGACGCAACGAGATGATCTCAAGCCAGATGCTTGAAGTTCTTGAAGAAACAG 9381
 QY 601 AGTAGTAACGAACCTGTTTGTTCGTGAGTTGCTTAAACCGGCTGCTTAAACAGATG 660
 DB 9380 AGTAGTAACGAACCTGTTTGTTCGTGAGTTGCTTAAACCGGCTGCTTAAACAGATG 9321
 QY 661 ATGAGCTCTGTATTCGGAAGAAGATGATGAGCTTGAAGAAACCAATGTTGAGTTACGTGA 720
 DB 9320 ATGAGCTCTGTATTCGGAAGAAGATGATGAGCTTGAAGAAACCAATGTTGAGTTACGTGA 9261
 QY 721 ATGCTCGAAGAGGTTATGATTTGCTCGAAAGCTTGAATGAGCTGATCACTTCTCTGG 780
 DB 9260 ATGCTCGAAGAGGTTATGATTTGCTCGAAAGCTTGAATGAGCTGATCACTTCTCTGG 9201
 QY 781 CTATCGAGGTTGATTCCTCAAGAGCTCGGCTGATGATTTGCACTCTGACCAAGAGTA 840
 DB 9200 CTATCGAGGTTGATTCCTCAAGAGCTCGGCTGATGATTTGCACTCTGACCAAGAGTA 9141
 QY 841 AACCGATTGTATCCCGGATTATATCGAACAACCGTAATCAACCGGATGATTGCTCTGCT 900
 DB 9140 AACCGATTGTATCCCGGATTATATCGAACAACCGTAATCAACCGGATGATTGCTCTGCT 9081
 QY 901 GATTTCGTGACGTTTCTCTCCCTCCATGAGTTCAAGTAATATCCGACCCGAGATA 960
 DB 9080 GATTTCGTGACGTTTCTCTCCCTCCATGAGTTCAAGTAATATCCGACCCGAGATA 9021
 QY 961 ATCCGCTTCTTGGG 976
 DB 9020 ATCCGCTTCTTGGG 9005

RESULT 2
 AC006418/c
 LOCUS
 DEFINITION
 complete sequence.
 AC006418
 AC006418
 AC006418.4 GI:20157771
 HTG.
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosid 1; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 84825)
 Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Shen,M., Vanhaken,S.E.,
 Barnstead,M.B., Mason,T.M., Bowman,C.L., Bowman,C.M.,
 Benito,M.-I., Carrera,A.J., Creasy,T.H., Buell,C.R., Town,C.D.,
 Niernman,W.C., Fraser,C.M. and Venter,J.C.
 Unpublished
 2 (bases 1 to 84825)
 Lin,X.
 Direct Submission
 Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 3 (bases 1 to 84825)
 Town,C.D. and Kaul,S.
 Direct Submission
 Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA, cdtown@igrr.org
 On Apr 18, 2002 this sequence version replaced gi:6598561.
 FEATURES
 source
 1..84825
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="2"
 /map="C106C03"
 /clone="F13A10"
 1..36153
 /note="overlap with BAC clone F11C10
 (AC006526:59064..95214)."
 1..36153
 /note="overlap with BAC clone F11C10
 (AC006526:59064..95416)."
 complement(1047..3029)
 /gene="At2g46480"
 /note="synonym: F13A10.1; predicted by gensecan"
 complement(join(<1047..1332,1409..1518,1595..1997,
 2074..2431,2503..2828,2916..3029))
 /gene="At2g46480"
 complement(join(1047..1332,1409..1518,1595..1997,
 2074..2431,2503..2828,2916..3029))
 /gene="At2g46480"
 /codon_start=1
 /product="hypotheetical protein"
 /protein_id="AAD20159.1"
 /db_xref="GI:4415929"
 /translation="MTDACCCKGNEDKRVPRFGHWITGKAFNDTPENLHRSJROEK
 RLEANEIMNDLSLOKLTAMARSRSDVAPLGNVITWKEVRRGKSFEDWLRMD
 OITMARVSGLAKFNINLALHOEIELOMLAMEBESDIDQEQVLDSIRMGQILA
 RAHOLYBECKVNTNRLAMLOQVEDELENSQYITPLQLASKALPDALHCTWELN
 EHLIDPILPMNPRENLENPLCYHVALPSDNVLAASVYVNSTVMAADPSHVEFLA
 TKLNFGASGSMFLINPREGATITVQREDPTWINSVSPVLSQESAMKKFEYKTA
 RBSVSGSENKIRYPRKYSMLNHLRTIPRIPEKLEKILFVDDDVVYQDLTEIMS
 IDLKGVENEDPKFCGWAVYGNVIFDLKWKKNNTETVHFQNLNENFTLMKLTLP
 PGLIFEVNLTQPLQKRWHLGLGIDYKGIIDVKIKERSAVIHVGNHKKPWTWGISKTOP
 YWTKTNFDPHPIFTCRLEPE"
 4060..4696
 /gene="At2g46490"
 /note="synonym: F13A10.2"
 4060..4696
 mRNA

CDS /gene="Ac2g46490"
 4100..4504
 /gene="Ac2g46490"
 /codon_start=1
 /product="unknown protein"
 /protein_id="AA20160.1"
 /db_xref="GI:4415930"
 /translation="MAADIFRSISFEGICSLDLSAIERPPYHNKGCALHDKSGACK
 NONRPPSCRRRSGSISIPRSMSEGIVAMMLFSSSSSSNLQSLSSSSLSLN
 LASLPADNATSEPRSSKQLRMTEDEDD"
 complement(4483..4517)
 /rpt_family="(Cn)n"
 /rpt_length(6756..6792)
 /rpt_family="At_rich"
 complement(11207..13602)
 /gene="Ac2g46500"
 /note="synonym: F13A10.3; similarity to ubiquitin family
 of proteins; supported by cDNA: gi_16930424_gb_Af419566
 1.Af419566"
 complement(join(11207..11764,11858..>13602))
 /gene="Ac2g46500"
 complement(join(11384..11764,11858..13177))
 /gene="Ac2g46500"
 /codon_start=1
 /product="expressed protein"
 /protein_id="AA20161.1"
 /db_xref="GI:4415931"
 /translation="MSSAGVALSPRSBPLIMPLVANSCLSYDDPTIMVLTPLPS
 VIMRVLPSISSEIVKTLPIQSYRGPNVKNQILGGRRLARSNNRPGYSEENILHS
 LVKLSDLYDLTKTTGGKCRFRVERGNIGVKQISKSGSDVDEDSILTSE
 KLEQSLINDICANDSVLHLVDRSAKAVKVEKNFELSLVAQADKKGRAKSI
 VPRKSLSEPVVNSKAVPLVVKDMQSAISDLKQNSPVASSEGTAAYFMQPGG
 KNEVGFKPIDEBPAENNPOGLPLSPNGSLKKGTVEGALREVAAYIIDHKSQN
 KSGVEIGIGAGVPPAMIECHLPGFNHPKGIKTKIGSLQMFENDSGCEDMGLSP
 VPEVHKSYSLDILRLANDRSGNITLMTKDSGLVAVPIDHGCLEPSEFCECTEMY
 MPOARKSYSLDEQVYRSLDAEDDILDKHGMMALPRLQRLISMLLKKGYERL
 TAEIGIGMRRLISKSLVEMVEMZEGANVLQGTSEALFLALSDVDVHDLSQEH
 "
 complement(14394..15192)
 /gene="Ac2g46505"
 /note="synonym: F13A10.20; supported by full length cDNA:
 Ceres: 11006"
 complement(<14394..>15192)
 /gene="Ac2g46505"
 complement(14656..15111)
 /gene="Ac2g46505"
 /codon_start=1
 /product="Expressed protein"
 /protein_id="AA15243.1"
 /db_xref="GI:20197780"
 /translation="MSLRITLIDLRORATLSKSLPFSMHSISASATAVNPGLG
 RDSLSPPAOLKDPSTNLTDRSTISSISGSELNKAKFSRSSRGYTNNSFLR
 KIPVPHIRHGKEILLADVQEMKTRNLIMSLGLQIIVLKDILFL"
 complement(15830..17530)
 /gene="Ac2g46510"
 /note="synonym: F13A10.4"
 complement(<15830..>17530)
 /gene="Ac2g46510"
 complement(15830..17530)
 /gene="Ac2g46510"
 /codon_start=1
 /product="putative bHLH transcription factor"
 /protein_id="AA20162.2"
 /db_xref="GI:20197775"
 /translation="NMMSDLWDDEDSVSAVGLFRANSNNQNLFTVWGT
 DTLTKSLSLVDVWPSSENFSMWYALFMQGTMSRSGQVAGMGDCCKREPEBESKV
 VRSNPFNNMGAEETWODMRKAVLOKHLRFGSDSDNYALERVTATELFLASAY
 FPFNHGCGAGRCYCSGSKIVMLSDAVNSSEYCFRPFMAKSGITIVWPLTDGVLIE
 LGSVYMLPEHIGFVQVQLFWRYTQPVWVTSNTMTGIIHLGGDLGAHAYPKKE
 LEVRNMLDEFTPOSMEGYNKKKPTREYVPOBDVYVLENNYVNNYKTOITFA
 GSAVLAASNSLITQOKESCTEKRPSSILAGAGVSYVDKRRPKRGKPAWGEE
 PLNHWERORRRKLNQRYALRSVYPLSMKSLSDAISTIKELQEVKTEBDE
 RVGDSLSLSNSNITTVESPEVDIQANEEVYVAVLSPDSHPARILLQANENSVL

	repeat_region	repeat_region	gene	mRNA	CDS
	17920..17946	/rpt_family="GA"n"	complement(18392..18454)		
	repeat_region	/rpt_family="At_F1ich"	21510..24428		
	gene	/gene="At2g46520"	/note="synonym: F13A10.5"		
		<21510..24428	/gene="At2g46520"		
		21510..24428	/gene="At2g46520"		
		/codon_start=1	/product="putative cellular apoptosis susceptibility protein"		
		/protein_id="A020163.1"	/db_xref="GI:4415933"		
		/translation="MEMNRETLVFLSGCFNLNTLSPIPEPRFTRALRSAPAAQAVGLAVLRALPAIDPCQTRAAAVNPKRLSGRPMADGSGI.SPVDSKEDKILTVLSIMLSASPIQSLSEALVYTGKDPKXKAPPLLELIANTQMAALAGYVSYNIGLSTASSTRKSCCTTIFSYLTFQDLPEFEDHKMEGGEFKYLSLNPALSTBESGSPFLKLEFSAACENINHYIEKNEEFPQGFNEFASVAVTLLRDVSKSPSDQLATTAIKELVSTSVHAI.PAGDNVAKELICOSIVIPNLSRVDESEIEMNATIEIRDMESSDVTRRIACGLKGLATNKTQVTEVLSLEIQRLSSFSANPSANKDCAIYLVASLTKAGGASVTRDLIDVONFANITLPELOSDVNSFPMKXGSLKFLTPMSPI.EKPFAMOLFPRVRLKASNVNHYSHVSGASTCEKILVBEAGNGRYAAGLSPIQLQMLNDALKPFSESENOYLMKCTIRNRYGVADISAEVAGPCIGVLTLSVCCNKPRTNNHYLFSSVAVLRACERDILSLAFETSLPFSLOMTIANDTETPLPGFQIQAQVYLNRPFLSENWMOIFLLILSPESWRKSGNVPAVRLQAFLOQAPHEVQENRLSGVLIETVLASSTDSQGYIILNTIENLDYSVIAFPYKGVWSALFTFQVOKRTVKRQKSLVETSLFLVHQCAQVLYVEMNTVQPNIIITAIYEHVHIEPMLKTIIMGSEVKTAVAATSLICEPTALIDPSAKLGMGLMDSIVTVLSVRPEOEVLDEPEMPEISENVGYTAAVAKLHNGKKERDPLNDKDPQFLVAVSARLSSASPRGYPIQIGENILNQANOTALIQCNANVCIA"			
	repeat_region	25088..25183	/rpt_family="Rf:gi 1100803 emb X93607.1 AT74A A.challiana microsatellite repeat region (clone 74A) #Rf: ATR0046 X93607 74A"		
	repeat_region	25090..25183	/rpt_family="(GA)n"		
	repeat_region	25090..25166	/rpt_family="Rf:gi 1100803 emb X93607.1 AT74A A.challiana microsatellite repeat region (clone 74A) #Rf: ATR0046 X93607 74A"		
	repeat_region	25090..25162	/rpt_family="Rf:gi 1100803 emb X93607.1 AT74A A.challiana microsatellite repeat region (clone 74A) #Rf: ATR0046 X93607 74A"		
	repeat_region	25090..25158	/rpt_family="Rf:gi 1100803 emb X93607.1 AT74A A.challiana microsatellite repeat region (clone 74A) #Rf: ATR0046 X93607 74A"		
	repeat_region	25090..25154	/rpt_family="Rf:gi 1100803 emb X93607.1 AT74A A.challiana microsatellite repeat region (clone 74A) #Rf: ATR0046 X93607 74A"		
	Query Match	61.3%; Score 976; DB 8; Length 84825;			
	Best Local Similarity	100.0%; Pred. No. 0;			
	Matches	976; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	ATGCGTACGAACTCGAAGCTCTTATCTTTGGCCCTTTTGTCCAAATGCGAGCTTTCTA	60		
DB	80058	ATGCGTACGAACTCGAAGCTCTTATCTTTGGCCCTTTTGTCCAAATGCGAGCTTTCTA	79999		
QY	61	AGCGAAACCAACTGGCTCTCCCTCTGCGCGGTACACATCATCTGCGTCGCGCATATCT	120		
DB	79998	AGCGAAACCAACTGGCTCTCCCTCTGCGCGGTACACATCATCTGCGTCGCGCATATCT	79933		
QY	121	CTCTCTTATGACATATCCCGGTGACCTGCTTGGGGGAAATACCTCTTCCGCGGCTTA	180		
DB	79938	CTCTCTTATGACATATCCCGGTGACCTGCTTGGGGGAAATACCTCTTCCGCGGCTTA	79877		

QY 181 ATATCGGTTTATACAAAACCGGAAAGTTATCCCGTCCAAAAGCTTCCCTTTGGTT 240
 Db 79878 ATATCGGTTTATACAAAACCGGAAAGTTATCCCGTCCAAAAGCTTCCCTTTGGTT 79819
 QY 241 GGAAGATGTCATCTATGTCAGACCTCTAGCTCACCGAGCAATGCTGATGAGCTGAG 300
 Db 79818 GGAAGATGTCATCTATGTCAGACCTCTAGCTCACCGAGCAATGCTGATGAGCTGAG 79759
 QY 301 AAATTCGAGCCAAAGAGCTCATGCGTTTCAGCTTGAAGAGACTGCGCTGATCGTACG 360
 Db 79758 AAATTCGAGCCAAAGAGCTCATGCGTTTCAGCTTGAAGAGACTGCGCTGATCGTACG 79699
 QY 361 TGCATTCGCAAGTACGAGAAAGATTCGAAATAGCCCGGTTTTCGATGAGACCGGTT 420
 Db 79698 TGCATTCGCAAGTACGAGAAAGATTCGAAATAGCCCGGTTTTCGATGAGACCGGTT 79639
 QY 421 AAAGATTCGCTTACTACTGATGATGTTTAAACAGCAATGCTTTCGACCAACGCTGTT 480
 Db 79638 AAAGATTCGCTTACTACTGATGATGTTTAAACAGCAATGCTTTCGACCAACGCTGTT 79579
 QY 481 TACTGCGGAAAGCTTCCCGTATGCTTGAACCACTCTCTTGTAGTACAAACAAATCAGA 540
 Db 79578 TACTGCGGAAAGCTTCCCGTATGCTTGAACCACTCTCTTGTAGTACAAACAAATCAGA 79519
 QY 541 AGAGCCGAGACGCAACGAGAGATCTCAAGCCAGATGCTTGAAGTTTCTTGAAGAAACAG 600
 Db 79518 AGAGCCGAGACGCAACGAGAGATCTCAAGCCAGATGCTTGAAGTTTCTTGAAGAAACAG 79459
 QY 601 AGTAGTACGAAACCTCTGTTTGTTCGTGAGTTGTTTAAACGCGCTGCTTAAACAATG 660
 Db 79458 AGTAGTACGAAACCTCTGTTTGTTCGTGAGTTGTTTAAACGCGCTGCTTAAACAATG 79399
 QY 661 ATGAGCTCTGATTCGCAACAGATGATGAGCTTGAAGAAACCAATGTTGAGTACGAGAA 720
 Db 79398 ATGAGCTCTGATTCGCAACAGATGATGAGCTTGAAGAAACCAATGTTGAGTACGAGAA 79339
 QY 721 ATGCTCGAAGAGTTATGATTTGCTCGAAAGCTTGAATGATGATGATCACTCTCTTGG 780
 Db 79338 ATGCTCGAAGAGTTATGATTTGCTCGAAAGCTTGAATGATGATGATCACTCTCTTGG 79279
 QY 781 CTATCGAGTTGATTCCTCGAAGACTCGGCTGATGATGATGATGATGATGATGATGATGAT 840
 Db 79278 CTATCGAGTTGATTCCTCGAAGACTCGGCTGATGATGATGATGATGATGATGATGATGAT 79219
 QY 841 AACCGTTTGTATCCCGGATTAATCCGAAACCGCTATCAACCGCTGATTTGCTCGT 900
 Db 79218 AACCGTTTGTATCCCGGATTAATCCGAAACCGCTATCAACCGCTGATTTGCTCGT 79159
 QY 901 GATTTCTGCAAGTTTGTCTCTCTCTCATGATGATGATGATGATGATGATGATGATGATGAT 960
 Db 79158 GATTTCTGCAAGTTTGTCTCTCTCTCATGATGATGATGATGATGATGATGATGATGATGAT 79099
 QY 961 ATCGCGCTTCTTTGGG 976
 Db 79098 ATCGCGCTTCTTTGGG 79083

RESULT 3
 ATH552847/c 740 bp DNA linear PLN 29-MAR-2003
 LOCUS Arabidopsis thaliana T-DNA flanking sequence, left border, clone 345B11.
 DEFINITION
 ACCESSION AJ552847
 VERSION AJ552847.1 GI:29368998
 KEYWORDS left border; T-DNA flanking sequence.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE 1 Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.,

Chauvin, S., Bechtold, N., Gruaud, C., DeRose, R., Pelletier, G., Lepoint, L., Caboche, M., and Lecharny, A.
 T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites
 EMBO Rep. 3 (12), 1152-1157 (2002)
 MEDLINE 22363535
 PUBMED 12446565
 REFERENCE 2 (bases 1 to 740)
 AUTHORS Balzergue, S.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE
 COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).
 FEATURES
 source
 1..740
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Wassilewskija"
 /db_xref="taxon:3702"
 /clone="345B11"
 /clone_1b="Arabidopsis thaliana T-DNA insertion lines"
 misc_feature
 1..740
 /note="T-DNA flanking sequence
 left border"
 BASE COUNT 192 a 172 c 197 g 179 t
 ORIGIN
 Query Match 14.6% Score 232; DB 8; Length 740;
 Best Local Similarity 99.4% Pred. No. 5e-121;
 Matches 332; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 200 CCGAAACGTTATTCCTCCGTCGAAAGCTTCCCTTGTGGAAGATGTCATCATGT 259
 Db 429 CCGAAACGTTATTCCTCCGTCGAAAGCTTCCCTTGTGGAAGATGTCATCATGT 370
 QY 260 CAAGCACTTACTCAACGAGCAATCGCTGATGAGCAATGCGAAGTGGAGGCAAGAGGC 319
 Db 369 CAAGCACTTACTCAACGAGCAATCGCTGATGAGCAATGCGAATTCGAGCAAGAGGC 310
 QY 320 TCATGCTTCAAGCTTGAAGAGACTCGGATGATGATGATGATGATGATGATGATGATGAT 379
 Db 309 TCATGCTTCAAGCTTGAAGAGACTCGGATGATGATGATGATGATGATGATGATGATGAT 250
 QY 380 AAGAGATTGTAATAGCCCGGTTTGTGCTGATGAGCAACCGCTTAAAGATCGCTTACTAC 439
 Db 249 AAGAGATTGTAATAGCCCGGTTTGTGCTGATGAGCAACCGCTTAAAGATCGCTTACTAC 190
 QY 440 TGAATTTAAGAGAGATTTGTTTGCACACAGAGTGTACTGCGAAGCTTGGCC 499
 Db 189 TGAATTTAAGAGAGATTTGTTTGCACACAGAGTGTACTGCGAAGCTTGGCC 130
 QY 500 GATTCGCTTGAACCATCTCTTTAGTACAAACA 533
 Db 129 GATTCGCTTGAACCATCTCTTTAGTACAAACA 96

RESULT 4
 ATH553770/c 504 bp DNA linear PLN 29-MAR-2003
 LOCUS Arabidopsis thaliana T-DNA flanking sequence, left border, clone 371B10.
 DEFINITION
 ACCESSION AJ553770
 VERSION AJ553770.1 GI:29370237

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 00:04:29 ; Search time 3521 Seconds
(without alignments)

10996.033 Million cell updates/sec

Title: US-10-022-025A-2

Perfect score: 1593

Sequence: 1 atgctacgaactcgaag.....tgcgcgtagcgagctaa 1593

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

BST:
1: em_estba:*
2: em_estlm:*
3: em_estlm:*
4: em_estlm:*
5: em_estlm:*
6: em_estlm:*
7: em_estlm:*
8: em_estlm:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est3:*
13: gb_est4:*
14: gb_est4:*
15: em_estlm:*
16: em_estlm:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	566.4	35.6	712	29	BZ428523 BONHW74TF
2	527	33.1	784	28	BH492184 BOHFJ05TR
3	495.8	31.1	809	29	BZ428529 BONHW74TR
4	476	29.9	674	28	BH451907 BOGSO44TR

C	5	428	26.9	745	28	BH510149
C	6	421.8	26.5	727	28	BH527239
C	7	399.2	25.1	659	29	BZ466758
	8	391.8	24.6	776	29	BZ443616
	9	330.8	20.8	705	28	BH977803
	10	330.8	20.8	732	28	BH534160
	11	329.4	20.7	826	14	CB344213
	12	322.6	20.3	638	14	CB343993
	13	314.8	19.8	720	10	BG648529
	14	299.6	18.8	576	12	BH451899
	15	298.6	18.7	791	12	BG39564
	16	297.2	18.7	804	28	BH431683
	17	286.2	18.0	612	9	AV826713
	18	285	17.9	765	28	BH581933
	19	283.4	17.8	764	28	BH652028
	20	271.2	17.0	455	9	A1995549
	21	268.4	16.8	827	14	CB655569
	22	265.4	16.7	764	13	BO403090
	23	262.2	16.5	846	13	BG518647
	24	251.2	15.8	540	12	BH250286
	25	248.8	15.6	549	28	B1973185
	26	247.2	15.5	739	28	BH547739
	27	242	15.2	616	14	CA030582
	28	238.8	15.0	691	9	AW132420
	29	238	14.9	740	14	CB344274
	30	236.6	14.9	540	29	BZ436712
	31	233.2	14.6	489	10	BZ345914
	32	231.2	14.5	641	12	BM063648
	33	229.6	14.4	635	14	CB628186
	34	229.2	14.4	656	9	AW622577
	35	224.4	14.1	835	10	BG447296
	36	220.4	13.8	454	10	BZ210324
	37	220.4	13.8	605	12	BZ226188
	38	217.6	13.7	810	10	BG583985
	39	213.6	13.4	712	14	CA783169
	40	212.6	13.3	468	9	BF586022
	41	212	13.3	653	10	BF595944
	42	211.8	13.3	495	10	BG580466
	43	211.8	13.3	537	9	AL365677
	44	207.6	13.0	531	10	BZ193362
	45	207.4	13.0	524	10	BZ998607

ALIGNMENTS

RESULT 1
BZ428523
LOCUS
DEFINITION
BONHW74TF BO.1.6.2 XB. loc Brassica oleracea genomic survey sequence.
ACCESSION
BZ428523
VERSION
BZ428523.1 GI:26671493
KEYWORDS
SOURCE
ORGANISM
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
Town, C.D., Van Aken, S., Uteirbeck, T., Koo, H. and Fraser, C.M.
AUTHORS
Whole genome shotgun sequencing of Brassica oleracea
TITLE
Unpublished
JOURNAL
Other GSSs: BONHW74TR
COMMENT
Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.

QY 541 AGAGCCGAGACGCAACAGAGATCTCAAGCAGATGTTGATTTCTTGAACACG 600
 DB 205 CGGCGGAAACGACAGAGCGAGTATCGCTAGCCAGATGATGAGGTTGCTCCAGAAACAG 146
 QY 601 ---AGATGATACGACACCTCTTTTGTTCGATGCTTAAACCGCGCTTAACAC 657
 DB 145 AGCAGACTACCGGCTCTGTGTTGTCGAGCTCTCAAGCGCGCTTACGAC 86
 QY 658 ATGATGCTCTGTATTCGACCAAGATATGAGCTTGAACCAACATGTTAGCT 717
 DB 85 ATGATGCTCTGTGTTTTCGACCAAGATATGAGCTTGAACCAACATGTTAGCT 26
 QY 718 GAAATGCTCGAAGAGTT 736
 DB 25 GAAATGCTCGAAGAGTT 7

RESULT 3
 BZ428529/c 809 bp DNA linear GSS 13-DEC-2002
 LOCUS BONHW74TR BO.1.6_2_KB_tot Brassica oleracea genomic clone BONHW74,
 DEFINITION genomic survey sequence.
 ACCESSION BZ428529
 VERSION BZ428529.1 GI:26671510
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 809)
 Town,C.D., Van Aken,S., Uterback,T., Koo,H. and Fraser,C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 Unpublished
 Other_GSSs: BONHW74TF
 COMMENT Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..809
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO100DH3"
 /db_xref="taxon:3712"
 /clone="BONHW74"
 /clone_1lb="BO.1.6_2_KB_tot"
 /note="Vector: pHOSt; Site 1: BstXI; 1.6-2 kb sheared
 total DNA inserted into pHOSt using BstXI linkers"
 BASE COUNT 241 a 216 c 146 g 206 t

Query Match 31.1%; Score 495.8; DB 29; Length 809;
 Best Local Similarity 87.6%; Pred.No. 1.1e-144;
 Matches 542; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 975 GGAGATGATATTCAGAGACAGACAGTGGCGGTCTTAATCGAGTGCCTCGCTAG 1034
 DB 742 GGAGATGATATTCAGAGACAGACAGTGGCGGTCTTAATCGAGTGCCTCGCTAG 683
 QY 1035 GATGCTCTTCATCCAGATATGCAATCAACGCTAACAAACAGCTGATCAAGTAGTCGG 1094
 DB 682 GATAGTGTTCATCCAGATATTCATTCAGACGATTCAGAGTGAACCTTGATGTTGTCGG 623
 QY 1095 GAAATCAGAGCCCTAGATGATCTGACTTGGCTTCACTTCATATCTGAACGGCTGAGT 1154
 DB 622 AAAATCAAGAGCTGATGATGCTGACTTGGCTTCACTTCATATCTGAACGGCTGAGT 563

QY 1155 GAAAGATATTTAGGCTTGATCCTCCAGGCCCACTTTCATGATGGGCGGTTGGCAT 1214
 DB 562 AAAAGAGTCTTGAAATGATCCACAGGCCCACTTTCATGATGGGCGGTTGGCAT 503
 QY 1215 AACAGACGATCGTTGATGCTCTTGTTCGCGACGAGACACAGCAATGCTGAACAT 1274
 DB 502 AACAGATGATCGTTGATGCTCTTGTTCGCGACGAGACACAGCAATGCTGAACAT 443
 QY 1275 GTGGCCGCTATGATGATCCACGCTGTGGGTTGATCCTTTGAGATTAACTGAGAG 1334
 DB 442 GTGGCAATAGCGATGATCCACGCTGTGGGTTGATCCTTTGAGATTAACTGAGAG 383
 QY 1335 GTTCGTGCAAAAGAGGTGAGTGGAGTTTCCGTTCTTGAGTCCGATTTGACACTGC 1394
 DB 382 ATTCTGACGAAAGAGTGAAGTGAAGTTCTCGGTTCTTGAGTCCGATTTGACACTGAC 323
 QY 1395 ACCTTTCGGGTGCGGTGCTCGAATTTGCCCGGGAAGATTTGCTTTACTACCGTTAT 1454
 DB 322 GCCTTCGGGTCTGCTGATGATCTGCTCGGGAAGATTTGATGACACCGTTAC 263
 QY 1455 GTTTTGAACGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1514
 DB 262 GTTCTGACCGCGACGCTGTTACAGATTTGAATGGGAGTGTCAAGTGTAAACGGGT 203
 QY 1515 TGACTTATCTGAGAAAGTGAAGCTTTCTTTCGAGATGGCTAATCTCTTCTGCTAATT 1574
 DB 202 TGACTTATCTGAGAAAGTGAAGCTTTCTTTCGAGATGGCTAATCTCTTCTGCTAATT 143
 QY 1575 GGGCGGTAGCGGCACTTAA 1593
 DB 142 GGGCGGTAGCGTACTTAA 124

RESULT 4
 BH451907/c 674 bp DNA linear GSS 12-DEC-2001
 LOCUS BOGS044TR BOGS Brassica oleracea genomic clone BOGS044, genomic
 DEFINITION survey sequence.
 ACCESSION BH451907
 VERSION BH451907.1 GI:17637618
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 674)
 Town,C.D., Van Aken,S., Uterback,T., Koo,H. and Fraser,C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 Unpublished
 Other_GSSs: BOGS044TR
 COMMENT Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..674
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO100DH3"
 /db_xref="taxon:3712"
 /clone="BOGS044"
 /clone_1lb="BOGS"
 /note="Vector: pHOSt; Site 1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pHOSt using BstXI linkers"
 BASE COUNT 191 a 194 c 130 g 159 t

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

FEATURES
source

1..727
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO100DH3"
/db_xref="taxon:3712"
/clone="BOHB106"
/clone_lib="BOHB"
/note="Vector: PHOS1, Site 1: BstXI, 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 197 a 204 c 145 g 181 t

ORIGIN

Query Match 26.5%; Score 421.8; DB 28; Length 727;

Best Local Similarity 81.1%; Pred. No. 2.5e-121; Indels 6; Gaps 1;

Matches 505; Conservative 0; Mismatches 112;

974 GGGAGATGATATTCAGAGAAACAGACACAGTGGCTTATTCAGATGATCTGCTA 1033
DB 644 GGGAAATGATATTTAGAGAACTGACACGGTGGCTGATTCAGATGATCTGCTA 585
QY 1034 GATGATGCTTCATCCAGATATGCAATCAGAGTACAAAGAGCTGATCAGTATG 1093
DB 584 GATGATGCTTCATCAGATATCTCAGACACAGTTCACAGAGCTTACAGAGTCTG 525
QY 1094 GGAATCAAGAGCCCTAGATGAATCTGACTTGGCTTCACTTCATATCTTAACGGCTG 1153
DB 524 GAGATCAAGAGCCCTAGATGAATCTGACTTGGCTTCACTTCATATCTGACGGCTG 465
QY 1154 TGAAGAAGTATGAGGCTTCATCTCAGAGCCCACTTCATATGAGGCGGCTTGCA 1213
DB 464 TCAAGAGATCTTGAAGCTTACCCGCAAGCTTCATCTGCTGAGGCGGCTTGCA 405
QY 1214 TAAACAGACAGATCTGATGATGCTTGTTCGAGAGGACCAACAGATGATG 1273
DB 404 TCAAGAGATCTTGAAGCTTACCCGCAAGCTTCATCTGCTGAGGCGGCTTGCA 345
QY 1274 TGTGGGCGCTATCGAGATATCAGAGTGGGCTGATCTCTTGGATTAACCTG 1333
DB 344 TGTGGGCTATAGCGAGATATCAGAGTGGGCTGATCTCTTGGATTAACCTG 285
QY 1334 GATTCGTGCAAAAGAGTGAAGTGGATTTTCGCTTGGCTGAGATTTGAAGCTT 1393
DB 284 GATTCGTGCAAAAGAGTGAAGTGGATTTTCGCTTGGCTGAGATTTGAAGCTT 225
QY 1394 CACCTTGGGCTGAGGCTGCTGAGATTTGGCCCGGAGAGATCTTGGATTAACCTG 1453
DB 224 CACCTTGGGCTGAGGCTGCTGAGATTTGGCCCGGAGAGATCTTGGATTAACCTG 165
QY 1454 TGTTCGTGCAAAAGAGTGAAGTGGATTTTCGCTTGGCTGAGATTTGAAGCTT 1507
DB 164 GATTCGTGCAAAAGAGTGAAGTGGATTTTCGCTTGGCTGAGATTTGAAGCTT 105
QY 1508 ACGGCTGATCTTATCTGAGAACTGAGGCTTTCTTGGAGATGAGGCTTACCTCTG 1567
DB 104 AGACCTGATCTTATCTGAGAACTGAGGCTTTCTTGGAGATGAGGCTTACCTCTG 45
QY 1568 CTAAATTCGCGCTGAGGCGCAGT 1590
DB 44 TTAAGTGTGCTGAGGCGCAGT 22

RESULT 7
B2466758 659 bp DNA linear GSS 13-DEC-2002
LOCUS B2466758
DEFINITION BONND11TR BO_1.6_2_KB_tot Brassica oleracea genomic clone BONND11,

ACCESSION
B2466758
VERSION
B2466758.1 GI:26759933
KEYWORDS
GSS.
SOURCE
ORGANISM

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

/ eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 659)

Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished

Other GSSs: BONND11TF

Contact: Chris Town

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

FEATURES
source

1..659
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO100DH3"
/db_xref="taxon:3712"
/clone="BONND11"
/clone_lib="BO_1.6_2_KB_tot"
/note="Vector: PHOS1, Site 1: BstXI, 1.6-2 kb sheared total DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 172 a 143 c 195 g 149 t

ORIGIN

Query Match 25.1%; Score 399.2; DB 29; Length 659;

Best Local Similarity 83.6%; Pred. No. 3.2e-114; Indels 11; Gaps 2;

Matches 480; Conservative 0; Mismatches 83;

3 GCGTACGAACTCGAAAGC--TCCTTAATCTTGGCCCTTTGTCGAATGACGCTTCTA 60
DB 565 GCGTACGAACTCGAAAGCTTTCCTGATCTTGGCCCTTTGTCGAATGACGCTTCTA 506
QY 61 AGCCAAACCAACCTGCTCTCCCTCCGCGCTGACATGATGAGTCCATATCT 120
DB 505 AGCCAAACCAACCTGCTCTCCCTCCGCGCTGACATGATGAGTCCATATCT 446
QY 121 CTCTTATGACCTATCCCGGTGACCTGCTTGGGGAATTAACCTTCCGCGGTTA 180
DB 445 CTCTTATGACCTATCCCGGTGACCTGCTTGGGGAATTAACCTTCCGCGGTTA 386
QY 181 ATATCCGCTTCAACAAACCGGAAAGTATTCGCGGCTTCCGCTTCCGCTT 240
DB 385 A-----CTAACAAACCGGAAACGCTTATTCGCGGCTTCCGCTTCCGCTT 335
QY 241 GGAAGATGTCATCTATGTCAGCACTTATGCTACCGACGAGATGCTGATGAG 300
DB 334 GGAAGATGTCATCTATGTCAGCACTTATGCTACCGACGAGATGCTGATGAG 275
QY 301 AAATTCGAGACCAAGAGGCTGATGAGTTCAGTTAGAGAGACTGAGTATGCTCAG 360
DB 274 AGATTCGAGACCAAGAGGCTGATGAGTTCAGTTAGAGAGACTGAGTATGCTCAG 215
QY 361 TGAATCCGACGATGAGCAAGAGATTCGATATGCTGATGAGTATGAGTATGAG 420
DB 214 TGAATCCGACGATGAGCAAGAGATTCGATATGCTGATGAGTATGAGTATGAG 155
QY 421 AAAGATCGGCTTACTGATGATGATTTTAAACAGAGCAATGATTTTGAACACGAGT 480
DB 154 AAAGATCGGCTTACTGATGATGATTTTAAACAGAGCAATGATTTTGAACACGAGT 95
QY 481 TACTGCGAAGGCTTCCGCGCTATGCTTGAACCACTCTTATGATCAAAACCAATCAGA 540

Db 94 TACTGCGAAGCGTGGGAAGATCGCTTGGAACCATCTTTCAGTCTTAACAATCAAA 35

QY 541 AGAGCCGAGACGACGACGAGTATCTCAAGC 574

Db 34 CGCGCGAAGACGACGAGGAGTATCGCTAACC 1

RESULT 8
BZ443616 776 bp DNA linear GSS 13-DEC-2002
LOCUS BONF28TR BO.1.6.2 KB tot Brassica oleracea genomic clone BONF28,
DEFINITION genomic survey sequence.
ACCESSION BZ443616
VERSION BZ443616.1 GI:26703871
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots; rosids
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 776)
Town,C.D., Van Aken,S., Uteback,T., Koo,H. and Fraser,C.M.
REFERENCE Whole genome shotgun sequencing of Brassica oleracea
TITLE Unpublished
JOURNAL
AUTHORS
COMMENT Other GSSs: BONF28TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..776
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone_lib="BONF28"
/clone_lib="BO.1.6.2 KB tot"
/note="Vector: pHS1; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into pHS1 using BstXI linkers"

BASE COUNT 196 a 156 c 207 g 217 t

ORIGIN

Query Match 24.6%; Score 391.8; DB 29; Length 776;
Best Local Similarity 81.1%; Pred. No. 7,8e-112;
Matches 481; Conservative 0; Mismatches 107; Indels 5; Gaps 2;

QY 967 GTCTTTGGAGATGATATTCAGAGACAGACAGTGGCGCTTATCGAGGATC 1026

Db 184 GTTTTACAGGAATGATTTAGAGAACTGACAGCGTGGCTATTCAGAGGAT 243

QY 1027 CTGCTTGGATGCTCTTCAATCAATATGCAATCAAGCTTCAAAAGAGTGGATTA 1086

Db 244 CTGCTTGGATGCTCTTCAATCAATATGCAATCAAGCTTCAAAAGAGTGGATTA 303

QY 1087 GTAGTGGGAATCAAGAGCCCTAGATGATCTGATCTGCTTCAATTCATCAACG 1146

Db 304 GTGTTAGGAGATCAAGAGCCCTAGATGATCTGATCTGCTTCAATTCATCAACG 363

QY 1147 GCTGTGGTGAAGAATATTCAGAGCTTCAATCTCAAGAGCCCTTCAATTCAGAGCCCT 1206

Db 364 GCTGTGGTGAAGAATATTCAGAGCTTCAATCTCAAGAGCCCTTCAATTCAGAGCCCT 423

QY 1207 TTGGCCATTAACAGACGATCGTTGATGCTCTTCTTCCGCAAGGACCAAGCAATG 1266

Db 424 TTAGCAATCAACATGATGATCAATTCAGAGCCCTGCTGCTCGGCAAGGACCAATG 483

QY 1267 GTAAATGTGGGCGGATGATGATCAACAGTGTGGTTGATCTTTGAGTTTAA 1326

Db 484 GTTAACATGTGGGCTATGACCCATGATTCACACGTGGGAAATTCGTTGAGTTGAA 543

QY 1327 CCTGAGAGCTTCTGTCGCAAAAGAGTGAAGTGGATTTTCGTTCTTGGGCGATTTG 1386

Db 544 CCGGAAGGTTGTATGCCAAGAGAGTGAAGTGAAGTTTCGTTCTTGGGCGATTTG 603

QY 1387 AGACTTGAACCTTTTGGGTCGGGTCGTCGATTTTCCCGGGAAGATCTTGGTTTACT 1446

Db 604 AGCTTGAACCTTTTGGGTCGGGTCGTCGATTTTCCCGGGAAGATCTTGGTTTACT 663

QY 1447 ACCGTTATGTTTGAACCGGATGATGATTCATGATTTGAATG--GGACGCTCCGAT 1503

Db 664 ACCGTTATGTTTGAACCGGATGATGATTCATGATTTGAATG--GGACGCTCCGAT 723

QY 1504 GGTAAAG--CGTGAATTTATTCAGAAACTGAGGCTTTTGGGAGATGGCT 1554

Db 724 GATGACAGACCGTGAATCTTGTCCAGAAACTGAGGCTTCTGTGAGATGGCT 776

RESULT 9
BH977803 705 bp DNA linear GSS 02-OCT-2002
LOCUS odel4a11.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
DEFINITION sequence.
ACCESSION BH977803
VERSION BH977803.1 GI:23460806
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots; rosids
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 705)
Delehaunty,K., Fawell,G., Fulton,L., McCombie,W.R., Miner,T., Nash
M., Rabinowicz,P.D. and Wilson,R.K.
REFERENCE Whole genome shotgun reads from Brassica oleracea
TITLE Unpublished
JOURNAL
AUTHORS
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Plate: odel4 row: a column: 11
Seq primer: -21UPPOT forward
Class: shotgun
High quality sequence start: 18
High quality sequence stop: 519.
Location/Qualifiers
1..705
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/note="Vector: pORW13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea TO1000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

BASE COUNT 173 a 143 c 193 g 193 t 3 others

ORIGIN

Query Match 20.8%; Score 330.8; DB 28; Length 705;
Best Local Similarity 76.1%; Pred. No. 1.4e-92;
Matches 458; Conservative 0; Mismatches 139; Indels 5; Gaps 4;

QY 967 GTCTTTGGAGATGATATTCAGAGACAGACAGTGGCGCTTATCGAGGATC 1026

Db 103 GTTTTACAGGAATGATTTAGAGAACTGACAGCGTGGCTATTCAGAGGAT 162

QY 1027 CTGCTTGGATGCTCTTCAATCAATATGCAATCAAGCTTCAAAAGAGTGGATTA 1086

```
FEATURES
    source
    location/Qualifiers
    1..732
    /organism="Brassica oleracea"
    /mol_type="genomic DNA"
    /strain="TO1000DH3"
    /db_xref="taxon:3712"
    /clone="BORC235"
```

Query Match	20.8%	Score 330.8;	DB 28;	Length 732;
Best Local Similarity	75.2%;	Pred. No. 1.4e-92;		
Matches 445; Conservative	0;	Mismatches 132;	Indels 15;	Gaps 2;

QY	238	GTTCGAAAGCATGTCACATCGTCAAGTCAAGACACTGAGCTACCCGAGCAATTCGGTATGACGT	237
Db	141	GTCCGGAAGCATGTCCTCATGTCTCCACAGCTTTAGCTACCGCTGCATAGCCGACAGCCCG	200
QY	298	GAGAAATTCGAGACCAAGAGCTCATGGCTTTCACTTAGAGAGACTCGCGTATGTC	357
Db	201	GAGAAATTCGAGACCAAGCTTTAATGGCGTTTAGCTGGAGATTACTCCGATGATGTC	260
QY	358	ACGTGCAATCCCGACGTACGAAAGAGATTTCTGAATAGCCCGGTTTTTGGTATCGACCG	417
Db	261	ACGTGCAACCTGTATGATGCTAAAGAAATTCAAACAGTCCGGTTTTTCGTGACCGCTCG	320
QY	418	GTTTAAAGATTCGGTTACTCATGATGTTTAAACAGAGCAATTGGTTTTTGACACACAGCT	477
Db	321	GTTTAAAGATTCAGGTAATCCCTTAATTTAGCCGGGCTATCGGTTTCGCTCTTAAGCGC	380
QY	478	GTTTACTGGCGAAGCTTCGCCGTATGCTTCGAAACCATCTCTTTAGTACAAACAAATC	537
Db	381	GTTTACTGGCGAAGCTTTGAGGAAAAATGCGCTAGTCAATCTTTTCAACCCGAAACAAATC	440
QY	538	AGAAAGCCGAGACGCAACGACGATGATCTCAAGCCAGATGTTGATGTTCTTGAATAA	597
Db	441	AAAGCCTCTGAAACGCAAGACGCTGTATCCGCAATCAGATCCGAAAGTCTTCGAATAA	500
QY	598	CAGAGTAGTAACA-----ACCTGTTTTGTCGTAGTGGCTTAAACGCGCTCGCTT	651
Db	501	CAGAGCACCGGCAACGAAAGACTCTGTTTGCTGCTGATCTTGATCAAAACGGACTCGCTT	560
QY	652	AACCAACATGATGTGCTCTGATTTGCGACAAAGATGATGACTTGAAAAAACCATG----	706
Db	561	AATATACATGATGTGCTCTGTTTGGGAAAAAATACGAGCTTTGAGCATGAACAACGAA	620
QY	707	-----TTGAGTTACGTGAATAGTGTGAAAGAGGTTATGATTCCTCAAGATCCCGGCTCA	762
Db	621	GTGAATGAGCTTCGGTGTTGGTTCGAAAGAAAGTTATGATTAATCTCCGAACCTTAACGG	680
QY	763	ACTATATCACTTTCCTTGCTATACGAAAGTTATCTCCTCAAGATCCCGGCTCA	814
Db	681	ACCGATCATCTCCCGTGGCTGTGCGAAATATGATCTCTCAAGAAATCCGCTCTA	732

RESULT 11	CB344213	826 bp	mRNA	linear	EST 10-APR-2003
LOCUS	CB344213				
DEFINITION	CA48EN0001.1af G02 Cabernet Sauvignon Leaf-CA48EN vitis vinifera				
ACCESSION	CB344213				
VERSION	CB344213.2	GI:29784005			
KEYWORDS	EST;				
SOURCE	Vitis vinifera				
ORGANISM	Vitis vinifera				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids				
	; Vitaceae; Vitis.				
REFERENCES	1 (bases 1 to 826)				
AUTHORS	Goes da Silva, F., Lim, H., Tandolino, A., Baek, J., Jones, K., Walker				
TITLE	M.A. and Cook, D.R.				
JOURNAL	Transcriptional responses of Vitis vinifera to infection by the				
COMMENT	bacterial pathogen Xylella fastidiosa				
	unpublished				
	On Mar 14, 2003 this sequence version replaced gi:28965180.				
	Contact: Douglas Cook, FMD				
	CAES Genome Facility				

UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcoc@ucdavis.edu
Seq primer: ACGGTACCGACATATGCCC.
Location/Qualifiers
1. 826
/organism="Vitis vinifera"
/mol type="mRNA"
/culti var="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CA48EN001_1bf_902"
/sex="Hermaphrodite"
/dev stage="Berry stage I"
/lab host="DH5alpha"
/note="Organ: Berry; Vector: pDNR; Site_1: Sfil; Site_2: Sfil; CA48EN is a cDNA library of Cabernet Sauvignon leaves. The leaves were collected on July 25, 2001, in Napa Valley, California, and represent leaves in mid-season development. These leaves were asymptomatic and verified to be non-infected with the bacterial pathogen, Xylella fastidiosa, based on a diagnostic assay using PCR and Xylella-specific primer pairs. cDNA made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
5'-AGCAGTGTATCAACGACGATGCGCATTCAGCGCGG-3' and
5'-ATTCTAGAGCCGACGCGCGCATG-DT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

BASE COUNT 179 a 218 c 224 g 203 t 2 others
ORIGIN

Query Match 20.7%; Score 329.4; DB 14; Length 826;
Best Local Similarity 68.7%; Pred. No. 4,3e-92;
Matches 433; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

935 CAGATTAATTAATCCGACCGGACATATATCGCGCTCTTTGGAGATGATATTCAGAGAA 994
3 CCATTAAGTTGATTCGATTCGATATATCGCGCTCTTTGGAGATGATATTCAGAGAA 62
995 CAGACAGATGCGGCTTAATCGAGTGAATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1054
63 CGATACGTTGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 122
1055 TGCATCAACGTTACAAACGAGCTGATCAAGTATGATGATGATGATGATGATGATGATG 1114
123 TCCATCAACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 182
1115 AATCTGATGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1174
183 AATCCGATTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 242
1175 ATCTTCAGCCCACTTCTATCAATGAGCCGCTTTGGCCATTAACAGACGATGATGATG 1234
243 ACCACCGGCGCCACGCTCTCTGCGCGCCGCTTTATCAATCAACATATCAACGATGATG 302
1235 GTCGTCTTGTTCGCGGAGGACACAGCAATGATGATGATGATGATGATGATGATGATGATG 1294
303 GCATACGCTGCTTCCCGGCTCAACGATGATGATGATGATGATGATGATGATGATGATGATG 362
1295 CACACGTTGAGTTGATCTTTGAGTTTAACTGAGAGGTTGCTGCGCAAAAGAGTG 1354
363 CACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 422
1355 AGGTGAGTTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1414
423 ACGTGAATCTCTCTGTTGATCGATCTCAAGGTCGCGCATTCGCGGATGCGGCAAA 482
1415 GATTTGCGCGGAGAAATCTTGTGTTTAACTGATGATGATGATGATGATGATGATGATG 1474
483 GAACCTGCGCGGAGAAATCTTGTGTTTAACTGATGATGATGATGATGATGATGATGATG 542

QY 1475 TACATGATTTGAATGGGACCGTCCGATGATAGCGCGTTGACTTATTCAGAAATCA 1534
DB 543 TGCATGATGTTGAGTGGGTTCCATCCGACCCGAGGATCTATCCGAAATGCTGA 602
QY 1535 GGCTTTCTTGGAGATGAGCTATATCTTCTGCTGCTTAAATTTGCGCGTGGCGAGTTTA 1593
DB 603 GGCTATCTCTGCGAAATGCGACACCTCTGACTGTGAGGAGTGGCGCCAGCGCAATGA 661

RESULT 12
CB343993 638 bp mRNA linear EST 10-APR-2003
LOCUS CA48EN001_1bf_902 Cabernet Sauvignon leaf-CA48EN Vitis vinifera
DEFINITION cDNA clone CA48EN001_1bf_902 5', mRNA sequence.
ACCESSION CB343993
VERSION CB343993.1 GI:28964960
KEYWORDS EST
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; Vitaceae; Vitis.
1 (bases 1 to 638)
Goes de Silva, F., Lim, H., Iandolo, A., Baek, J., Jones, K., Walker
M.A. and Cook, D.R.
Transcriptional responses of Vitis vinifera to infection by the
bacterial pathogen Xylella fastidiosa
Unpublished
Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcoc@ucdavis.edu
Seq primer: ACGGTACCGACATATGCCC.
Location/Qualifiers
1. 638
/organism="Vitis vinifera"
/mol type="mRNA"
/culti var="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CA48EN001_1bf_902"
/sex="Hermaphrodite"
/dev stage="Berry stage I"
/lab host="DH5alpha"
/note="Organ: Berry; Vector: pDNR; Site_1: Sfil; Site_2: Sfil; CA48EN is a cDNA library of Cabernet Sauvignon leaves. The leaves were collected on July 25, 2001, in Napa Valley, California, and represent leaves in mid-season development. These leaves were asymptomatic and verified to be non-infected with the bacterial pathogen, Xylella fastidiosa, based on a diagnostic assay using PCR and Xylella-specific primer pairs. cDNA made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
5'-AGCAGTGTATCAACGACGATGCGCATTCAGCGCGG-3' and
5'-ATTCTAGAGCCGACGCGCGCATG-DT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

FEATURES
source
1. 638
/organism="Vitis vinifera"
/mol type="mRNA"
/culti var="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CA48EN001_1bf_902"
/sex="Hermaphrodite"
/dev stage="Berry stage I"
/lab host="DH5alpha"
/note="Organ: Berry; Vector: pDNR; Site_1: Sfil; Site_2: Sfil; CA48EN is a cDNA library of Cabernet Sauvignon leaves. The leaves were collected on July 25, 2001, in Napa Valley, California, and represent leaves in mid-season development. These leaves were asymptomatic and verified to be non-infected with the bacterial pathogen, Xylella fastidiosa, based on a diagnostic assay using PCR and Xylella-specific primer pairs. cDNA made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
5'-AGCAGTGTATCAACGACGATGCGCATTCAGCGCGG-3' and
5'-ATTCTAGAGCCGACGCGCGCATG-DT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

BASE COUNT 133 a 172 c 189 g 144 t
ORIGIN

Query Match 20.3%; Score 322.6; DB 14; Length 638;
Best Local Similarity 69.4%; Pred. No. 5.1e-90;
Matches 439; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 935 CAGATTAATTAATCCGACCGGACATATATCGCGCTCTTTGGAGATGATATTCAGAGAA 994
DB 3 CCATTAAGTTGATTCGATTAATATCGCGCTCTTTGGAGATGATATTCAGAGAA 62

QY 995 CAGACAGATTGCGGCTTAAATCGATGATCTCGTAGAGTGGTCTTCAATCCAGATA 1054
 Db 63 CGGATACGGTGGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 122
 QY 1055 TCCATCAACGATCAAAACGATGATCAAGTATGATGATGATGATGATGATGATGATG 1114
 Db 123 TCCATCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 182
 QY 1115 AATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1174
 Db 183 AATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 242
 QY 1175 AATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1234
 Db 243 AATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302
 QY 1235 GTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1294
 Db 303 GTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 362
 QY 1295 GTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1354
 Db 363 GTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
 QY 1355 AATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1414
 Db 423 AATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 482
 QY 1415 AATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1474
 Db 483 AATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 542
 QY 1475 AATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1534
 Db 543 AATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
 QY 1535 AATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1594
 Db 603 AATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 662
 RESULT 13
 B648529
 LOCUS EST10148 HOGA Medicago truncatula cDNA clone phOGA-2355 5' end,
 DEFINITION mRNA sequence.
 ACCESSION B648529
 VERSION B648529.1 GI:13783641
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosidis I; Fabales; Fabaceae; Papilionoideae; Trifoliales;
 Medicago.
 1 (bases 1 to 720)
 REFERENCE Hahn, M.G., Ojane-Reuhs, T., Samac, D., Town, C.D., Van Aken, S.,
 AUTHORS Uterback, T., Chen, J., and Fraser, C.M.
 TITLE ESTs from roots of Medicago truncatula treated with
 JOURNAL oligolacturonides of DP 6-20
 COMMENT Unpublished
 Contact: Michael G. Hahn
 Complex Carbohydrate Research Center
 University of Georgia
 220 Riverbend Road, Athens, GA 30602-4712, USA
 Tel: 706-542-4457
 Fax: 706-542-4412
 Email: hahn@ccrc.uga.edu
 G392135e TIGR sequence name: MTMCM27TK More information is
 available at: www.medicago.org
 Seq primer: Skmod (CTA GAA CTA gtc gat cc).
 FEATURES
 source 1..720
 location/Qualifiers

/organism="Medicago truncatula"
 /mol_type="mRNA"
 /cuiilar="A17"
 /db_xref="taxon:3880"
 /clone="phOGA-2355"
 /cues_type="3 day old seedling roots"
 /dev_stage="24 hours after treatment in the dark at 26 C
 with 0.5 mg/ml oligolacturonides (DP 6-20) in the
 presence of 100 ug/ml Gentamicin"
 /lab_host="XLOLR"
 /clone_11b="HOGA"
 /note="Vector: Bluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
 was directionally ligated into the UniTap XR vector from
 StrataGene and packaged using GigaPack III Gold packaging
 extracts. Plasmids containing cDNA inserts were excised
 from the recombinant lambda-Zap phage using Ex-aseist
 helper phage and propagated in SOLR cells."
 BASE COUNT 179 a 128 c 201 g 212 t
 ORIGIN
 Query Match 19.8%; Score 314.8; DB 10; Length 720;
 Best Local Similarity 66.7%; Pred. No. 1,6e-87;
 Matches 483; Conservative 0; Mismatches 232; Indels 9; Gaps 2;
 716 GTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 775
 Db 2 GTGGGTTGTTGATCAAGGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 61
 QY 776 GTGGGTTGTTGATCAAGGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 835
 Db 62 CATTTGAAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121
 QY 836 AGTTAAACCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 895
 Db 122 AAGTAAACCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 181
 QY 896 CTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 955
 Db 182 ACAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 241
 QY 956 ACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1015
 Db 242 ACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 301
 QY 1016 TCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1075
 Db 302 TAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
 QY 1076 AGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1132
 Db 362 AGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421
 QY 1133 TTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1192
 Db 422 TGGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 481
 QY 1193 TATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1252
 Db 482 TATCGTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 541
 QY 1253 GATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1312
 Db 542 GATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 601
 QY 1313 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1372
 Db 602 CACTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 655
 QY 1373 TTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1432
 Db 656 TTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 715

QY	1433	ATCT	1436	
Db	716	ATTT	719	
RESULT 14				
LOCUS	BH451899			
DEFINITION	BH451899	576 bp	DNA	linear
ACCESSION	BH451899			
VERSION	BH451899.1	GI:17637610		
KEYWORDS	GSS.			
SOURCE	Brassica oleracea			
ORGANISM	Brassica oleracea			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euposide II; Brassicales; Brassicaceae; Brassica.			
AUTHORS	1 (bases 1 to 576)			
TITLE	Town,C.D., Van Aken,S., Uterback,T., Koo,H. and Frazer,C.M.			
JOURNAL	Whole genome shotgun sequencing of Brassica oleracea			
COMMENT	Unpublished			
	Other GSSs: BOGSO44TR			
	Contact: Chris Town			
	TIGR			
	9712 Medical Center Drive, Rockville, MD 20850, USA.			
	Tel: 301-838-3523			
	Fax: 301-838-0208			
	Email: cdtown@tigr.org			
	DNA is from a doubled haploid provided by Tom Osborn.			
	Seq primer: TP			
	Class: sheared ends.			
FEATURES	Location/Qualifiers			
source	1..576			
	/organism="Brassica oleracea"			
	/mol_type="genomic DNA"			
	/strain="TO1000DH3"			
	/db_xref="taxon:3712"			
	/clone="BOGSO44"			
	/clone_1fb="BOGS"			
	/note="Vector: pHOSt, Site 1. BstXI, 2-3 kb sheared genomic DNA inserted into pHOSt using BstXI linkers"			
BASE COUNT	137 a 162 c 114 g 163 t			
ORIGIN				
Query Match	18.8%; Score 299.6; DB 28; Length 576;			
Best Local Similarity	83.9%; Pred. No. 8.7e-83;			
Matches 334; Conservative	0; Mismatches 59; Indels 9; Gaps 1			
QY	1	ATGGCTACGAACTCGAAGCTCTTAATCTTTTGCCCTTTTGTCGAAATGCAAGCTTCTA	60	
Db	164	ATGGCTACGAACTCGAAGCTCTTAATCTTTTGCCCTTTTGTCGAAATGCAAGCTTCTA	223	
QY	61	AGGCAACCAACCTGGCTCTTCCTCCCTCCGCGGTGACAAATCATCTGCGCATATCT	120	
Db	224	AGGCAACCAACCTGGCTCTTCCTCCCTTCCTTAATCTGCAATGCTATCTGCGCTTTCT	283	
QY	121	CTCTTTTATGACCTATCCCGGTGACCTGCTTGGGGAAATACCTCTTGGCGGTTA	180	
Db	284	CTCTTTTATGACCTATCCCGGTGACCTGCTTGGGGAAATACCTCTTGGCGGTTA	343	
QY	181	ATATCCGGTTCATACCAACCGGAAAGCTTATCCCGGTGACCAAAAGGCTCCCTTGTT	240	
Db	344	A-----CTAACAAACCGGAAAGCTTATCCCGGTGACCAAAAGGCTCCCTTGTT	394	
QY	241	GGAAGCATCTCACTCATGTCAAGACCTTATGCTGACCGAGCAATGCTGATGCACTGAG	3000	
Db	395	GGAAGCATCTCTCATGTCCAGGCTCTGACACACCAAGCAATGCTGATGCACTGAG	4544	
QY	301	AAATTGGAGCGCAAGAGGCTCATGGCTTACAGCTTAGAGAGATCGCGATGCTGACG	3600	
Db	455	AGATACGAGCGCAAGAGGCTCATGGCTTACAGCTTAGAGAGATCGCGATGCTGACG	5144	

QY	361	TGCAATCCGACGCTAGCGAAAGAGATTCTGAAATAGCCCGCTTTGGCTATGACGCGTT	420
Db	515	TGCAACCCCGACGCTGACTAAAGAGATTGGAATAGTCGAGTTTGGCTAGCCGTCGTT	574
QY	421	AA 422	
Db	575	AA 576	
RESULT 15			
	BG839564/c	791 bp	EST 25-MAY-2001
	LOCUS	Gm01_13h01	F Gm01_AAFc_ECORC Glycine max cold_stressed_leaves
	DEFINITION	Glycine max cDNA Clone Gm01_13h01, mRNA sequence.	
	ACCESSION	BG839564	
	VERSION	BG839564.1	GI:14205866
	KEYWORDS	EST.	
	SOURCE	Glycine max (soybean)	
	ORGANISM	Glycine max	
	REFERENCE	Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids	
	AUTHORS	Singh, J.A., Robert, L.S., Lu, B., Zhu, L., De Moor, A., Couroux, P., Harris, L.J., Hattori, D.I., Quillet, T., Sprott, D. and Tinker, N.A.	
	TITLE	Expressed Sequence Tags from Cold-Stressed Glycine max (soybean) Leaves	
	JOURNAL	Unpublished	
	COMMENT	Contact: Singh, J.A. Eastern Cereal and Oilseed Research Centre Agriculture and Agri-Food Canada KM Neelby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A 0C6, Canada Tel: (613) 759-1662 Fax: (613) 759-1701 Email: singhja@em.agr.ca.	
	FEATURES	Location/Qualifiers	
	source	1..791	
		/organism="Glycine max"	
		/mol_type="mRNA"	
		/culivar="Maple Arrow"	
		/db_xref="taxon:3847"	
		/clone="Gm01_13h01"	
		/tissue_type="Leaves"	
		/dev_stage="15-day seedlings"	
		/clone_11b="Gm01_AAFc_ECORC_glycine_max_cold_stressed_leav	
		8"	
		/note="Vector: Bluescript SK-/XhoI-EcoRI; Site 1: EcoRI; Site 2: XhoI; Plants were grown 12 days from seeds, treated at 20C for 3 days, then leaves were collected. Library was prepared using, Unitaq ZAP-cDNA synthesis kit /Packaged GigaPack III Gold."	
	BASE COUNT	202 a	243 c 179 g 166 t 1 others
	ORIGIN		
	Query Match	18.7%	Score 298.6; DB 12; Length 791;
	Best Local Similarity	69.3%	Pred. No. 2.2e-82;
	Matches 481; Conservative	0; Mismatches 200;	Indels 13; Gaps 5;
QY	896	CTCTGATTTCTGTCAGCTTTTGTCT-CTCCCTCCAGTGTTCAGATAATTATTCGAC-C	952
Db	758	CCGGGATTTGCTTCAAGTTTGTCTCTCTCCATCAAGTCCCGATAATTGTCTCACTT	699
QY	953	CGGACATTAATGCCCTTTCTTTGGGAGATGATTTCAAGGAAACAGACAGCTT-CCGCGTC	1011
Db	698	CCGACATGATTTCTCTCTCTCTGGGAAAGATATTATAGGGGAGCAGACCCGTTGGCGTT	639
QY	1012	TTAATCGAGTGAATCCTCGCTAGATGGTCTTTCATCCAGATATGCAATCAACGGTACAA	1071
Db	638	TTGATTGAGTGAATATCGGCGAGATGGTGTCTTCATCCGAGAGTGCAAGAGAAAGTTC	579
QY	1072	AACGAGCTGATCAAGATGATCGGGGAATCAAGAGCCCTAAGATGATCTGACTTGGCTTCA	1131

```
Db 578 GAGAGATTGACCGCGGTGTTAGG---GGTGGCGCTTTGACGGAGAGGTCTGTGGCGCG 522
QY 1132 CTTCATATCTAACGGCTGTGTGAAAGAGATATGAGGCTTCATCTCCAGGCCACTT 1191
Db 521 ACGCGATCTTGCGGGGTGTGAAGAGGTCTGAGGCTGCACCGCGGGCCGCTT 462
QY 1192 CTATCATGGGCCCCGTTGGCCATPACAGACAGATGTTGATGTGTCTGTCTCGGCA 1251
Db 461 CTCTCGGGCCCGCTGGCCATCACTGATPACAGACATGATGTATCACTGTGCTGCG 402
QY 1252 GGGACACAGCAATGTGAACATGTGGGCCGTATCCATGATCCACACGTGTGGTTGAT 1311
Db 401 GGGACACCGCTATGTATATGTGGGCCATAGCAAGGACCCGGAGGTGTGCTGAC 342
QY 1312 CCTTTGAGTTTAACTGTGAGGTTGTGTGCAAAAAGAGTGAAGTGTGTTTGGTT 1371
Db 341 CCACTTGAGTTCAAGCCCGAGAGGTTCATGGGTCTGGAA-----ACGAGTTTCTGTT 288
QY 1372 CTGGGTCGGATTGAGACTTGCACTTTGGGTCGGGTGTGTGGAATTTGCCCCGGGAAG 1431
Db 287 TTGGGGTCGATCTGAGACTCGCTCCATTGGTTGGGTGCGAAGACATGCCCGGGAAG 228
QY 1432 AATCTTGGTTTAACTACCGTTATGTGTTGACGCGCATGATGTTAATGATGATG 1491
Db 227 ACTTTGGGTTGAGCACCGTAACTTCTGGGTGGCTTGCTTTGCATGAGTTGAATG 168
QY 1492 GGAACCGTCGATGTGTAAAGGCGTTGACTTATCTGAGAACTGAGGCTTCTTGCGAGATG 1551
Db 167 CTACCGTCTGATGAGCCAAAGTTGATCTAACGAGGTGCTGAGGCTCTGTGTGAATG 108
QY 1552 GCTAATCTCTTCTCTGCTAAATTGCGCGGTAGGC 1585
Db 107 GCTAACCACCTCATTTGTTAAAGTTGCGCCTAGGC 74
```

Search completed: January 17, 2004, 04:15:21
Job time : 3524 secs

BZ380060
LOCUS BZ380060 262 bp DNA linear GSS 26-NOV-2002
DEFINITION SALK_114536.29.00.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_114536.29.00.x, genomic
survey sequence.
BZ380060
ACCESSION BZ380060.1 GI:25472537
KEYWORDS
SOURCE GSS
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 262)
Alonso, J.M., Leisner, T.J., Barajas, P., Chen, H., Cheuk, R., Gadrinab,
C., Jeske, A., Karnes, M., Kim, C.U., Parker, H., Prednis, L., Shinn, P.,
Zimmerman, J., and Ecker, J.R.
A sequence-indexed library of insertion mutations in the
Arabidopsis Genome
Unpublished
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated intron of At2g46660.
Class: TDNA tagged.
FEATURES
source
1..262
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_114536.29.00.x"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT 64 a 56 c 62 g 75 t 5 others
ORIGIN
Query Match 3.1%; Score 50; DB 29; Length 262;
Best Local Similarity 100.0%; Pred. No. 1.3e-16;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1112 ATGATCTGACTGCTGCTCACTTCATCTACGCTGTGTGTAAGAA 1161
Db 160 ATGATCTGACTGCTGCTCACTTCATCTACGCTGTGTGTAAGAA 209

COMMENT
Other_GSSs: BOGSO44TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..674
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGSO44"
/clone="BOGSO44"
/note="Vector: pROSL; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pROSL using BstXI linkers"
BASE COUNT 191 a 194 c 130 g 159 t
ORIGIN
Query Match 2.7%; Score 43; DB 28; Length 674;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1286 CGCATGATCCACACGTGTGCTTATCCTTTGAGTTTAAAC 1328
Db 410 CGCATGATCCACACGTGTGCTTATCCTTTGAGTTTAAAC 368

RESULT 4
BHS10149/c
LOCUS BHS10149 745 bp DNA linear GSS 13-DEC-2001
DEFINITION BOGTQ49TR BOGT Brassica oleracea genomic clone BOGTQ49, genomic
survey sequence.
BHS10149
ACCESSION BHS10149.1 GI:17718239
KEYWORDS
SOURCE GSS
ORGANISM Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 745)
Town, C.D., Van Aken, S., Uterback, T., Koo, H., and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Other_GSSs: BOGTQ49TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
source
1..745
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGTQ49"
/clone="BOGTQ49"
/note="Vector: pROSL; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pROSL using BstXI linkers"
BASE COUNT 219 a 186 c 137 g 203 t
ORIGIN
Query Match 2.7%; Score 43; DB 28; Length 745;

Tue Jan 20 17:43:21 2004

us-10-022-025a-2.oligo.rst

Page 3

Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1286 CGCATGATCCACAGTGGGTGATCCCTTGAGTAAACC 1328
DB 197 CGCATGATCCACAGTGGGTGATCCCTTGAGTAAACC 155

RESULT 5
AL946977 82 bp DNA linear GSS 24-OCT-2002
LOCUS Arabidopsis thaliana T-DNA flanking sequence GR-300D11-015600,
DEFINITION genomic survey sequence.
ACCESSION AL946977 GI:24403599
VERSION AL946977.1
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Seadler,H.
TITLE A pipeline for automated high-throughput generation of FSTs
(flanking sequence tags) from Arabidopsis thaliana T-DNA
transformed lines
JOURNAL Unpublished
2
Rosso,M., Strizhov,N., Li,Y., Reis,B., Dekker,K. and Weishaar,B.
A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
for flanking sequence tag based reverse genetics
JOURNAL Unpublished
3 (bases 1 to 82)
Li,Y., Strizhov,N., Rosso,M. and Weishaar,B.
Direct Submission
Submitted (21-OCT-2002) Weishaar B., Max-Planck-Institut fuer
Zuechtungsforshung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
This sequence is recovered from the left border of the T-DNA. It
indicates an insertion close to or within gene At2g46660. The
sequences are generated at the MPI for Plant Breeding Research in
the context of the GABI-Kat project. GABI-Kat is part of the German
Plant Genomics program designated 'GABI'. Information on line
availability can be found at:
http://www.mpiz-koeln.mpg.de/GABI-Kat/.

FEATURES

source
1..82
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GR-300D11-015600"
/note="Arabidopsis thaliana T-DNA insertion lines"
/note="PCR was performed on DNA from Arabidopsis thaliana
vector pACT1. The lines contain one or more T-DNA
insertions. The DNA fragment(s) resulting from the PCR
were directly sequenced to determine the genomic sequence
flanking the insertion. Sequences displaying significant
similarity to the A. thaliana nuclear genome sequence were
processed for submission. T-DNA derived sequences were
removed"

BASE COUNT 17 a 12 g 19 t 21 others
ORIGIN

Query Match 2.4%; Score 38; DB 29; Length 82;
Best Local Similarity 100.0%; Pred. No. 8e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 996 AGACACAGTTGGGTCTTAATCGAGTGATCTCGCTA 1033
DB 45 AGACACAGTTGGGTCTTAATCGAGTGATCTCGCTA 82

RESULT 6
BZ428529/c 809 bp DNA linear GSS 13-DEC-2002
LOCUS BONHW74TR BO.1.6.2 KB tot Brassica oleracea genomic clone BONHW74,
DEFINITION genomic survey sequence.
ACCESSION BZ428529 GI:26671510
VERSION BZ428529.1
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea

REFERENCE
AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished
COMMENT Other GSSs: BONHW74TF
Contact: Chris Town
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES

source
1..809
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO100DH3"
/db_xref="taxon:3712"
/clone="BONHW74"
/note="Vector: pROST; Site: 1: BstXI, 1.6-2 kb sheared
total DNA inserted into pROST using BstXI linkers"
BASE COUNT 241 a 216 c 146 g 206 t
ORIGIN

Query Match 2.3%; Score 37; DB 29; Length 809;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1244 TTCCGGCAGGACCAAGCATGTGTGAACATGTGGC 1280
DB 473 TTCCGGCAGGACCAAGCATGTGTGAACATGTGGC 437

RESULT 7
BZ428523 712 bp DNA linear GSS 13-DEC-2002
LOCUS BONHW74TF BO.1.6.2 KB tot Brassica oleracea genomic clone BONHW74,
DEFINITION genomic survey sequence.
ACCESSION BZ428523 GI:26671493
VERSION BZ428523.1
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea

REFERENCE
AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished
COMMENT Other GSSs: BONHW74TR
Contact: Chris Town
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208

Tue Jan 20 17:43:21 2004

us-10-022-025a-2.011go.rst

Page 4

Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

FEATURES

source

1..712
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOHRW74"
/clone_id="BO_1.6_2_KB_tot"
/note="Vector: PHOS1; Site_1: BstXI; 1.6-2 kb sheared
total DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT

168 a 179 c 177 g 188 t

ORIGIN

Query Match

2.2%; Score 35; DB 29; Length 712;
Best Local Similarity 100.0%; Pred. No. 6.5e-08;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

637 AAACGGCGCTGCTTAACATGATGCTCTGT 671

DB

387 AAACGGCGCTGCTTAACATGATGCTCTGT 421

RESULT 8

BH431683/c

BOHRT46TR BOHR Brassica oleracea genomic clone BOHRT46, genomic

LOCUS

survey sequence.

DEFINITION

BH431683

VERSION

KEYWORDS

SOURCE

ORGANISM

Brassica oleracea

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosid1; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 804)

Town, C.D., Van Aken, S., Uteperack, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-9523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..804
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOHRT46"
/clone_id="BOHR"
/note="Vector: PHOS1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT

241 a 182 c 155 g 226 t

ORIGIN

Query Match

2.0%; Score 32; DB 28; Length 804;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

835 AAGTAACCGGTTGTATCCGATTAATC 866

DB

545 AAGTAACCGGTTGTATCCGATTAATC 514

Search completed: January 17, 2004, 07:16:37
Job time : 3522 secs

Tue Jan 20 17:43:14 2004

us-10-022-025a-1.oligo.rng

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 13:41:32 : Search time 548 Seconds
(without alignments)
9896.297 Million cell updates/sec

Title: US-10-022-025a-1
Perfect score: 2009
Sequence: 1 aataataataataatgta.....atacagatttagatctt 2009
Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 30

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_19Jun03: *
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

No matches found

Search completed: January 16, 2004, 16:01:22
Job time : 548 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 15:41:01 ; Search time 123 Seconds
(without alignments)
7209.258 Million cell updates/sec

Title: US-10-022-025A-1
Perfect score: 2009
Sequence: 1 ataataataataaatgta.....atagagattagagatttt 2009

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 563978 seqs, 220691566 residues

Word size : 30

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents_NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:**
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:**
- 3: /cgn2_6/prodata/1/ina/5A_COMB.seq:**
- 4: /cgn2_6/prodata/1/ina/5B_COMB.seq:**
- 5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:**
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq:**

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description

No matches found

Search completed: January 16, 2004, 19:19:09
Job time : 123 secs

Tue Jan 20 17:43:15 2004

us-10-022-025a-1.oligo.rnpb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 18:05:12 ; Search time 662 Seconds
(without alignments)

10696.752 Million cell updates/sec

Title: US-10-022-025A-1

Sequence: 1 aataataataataatgta.....atagagaatgagatttt 2009

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2324096 seqs, 1762381658 residues

Word size : 30

Total number of hits satisfying chosen parameters: 3

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US05_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2009	100.0	2009	15 US-10-022-025A-1	Sequence 1, Appli
2	976	48.6	1593	10 US-09-349-385-5	Sequence 5, Appli
3	976	48.6	1593	15 US-10-022-025A-2	Sequence 2, Appli

ALIGNMENTS

TITLE OF INVENTION: Acid
FILE REFERENCE: 3310 0003
CURRENT APPLICATION NUMBER: US/10/022, 025A
CURRENT FILING DATE: 2003-01-09
PRIOR APPLICATION NUMBER: US 60/254, 819
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 2009
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: gene
LOCATION: (1)..(2009)
US-10-022-025A-1

Query Match 100.0%; Score 2009; DB 15; Length 2009;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2009; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AATATAATATAATATGATGCAATTTTGTATATAATAGATTGATTAATATAATA	60
DB	1	AATATAATATAATATAATGATGCAATTTTGTATATAATAGATTGATTAATATAATA	60
QY	61	ATACCAAGCCATGCTGCTCATCTTTAAGCCCAAGGCCAAGACACTCTCTTTTGT	120
DB	61	ATACCAAGCCATGCTGCTCATCTTTAAGCCCAAGGCCAAGACACTCTCTTTTGT	120
QY	121	TATCT	180
DB	121	TATCT	180
QY	181	CTCCCTATATCT	240
DB	181	CTCCCTATATCT	240
QY	241	CTCCCT	300
DB	241	CTCCCT	300
QY	301	CGGTGACCTGCT	360
DB	301	CGGTGACCTGCT	360
QY	361	CGGAAACGTTATTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	420
DB	361	CGGAAACGTTATTCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	420
QY	421	AAGCACTTACTCTACCAAGCAATCGCTGAAGCAAGCAAGCAAGCAAGCAAGCA	480
DB	421	AAGCACTTACTCTACCAAGCAATCGCTGAAGCAAGCAAGCAAGCAAGCAAGCA	480
QY	481	CATGCTTTCAGCTTACGAGAGACTCGCGTATCGTCACTGCAATCCGACGTAGCGAA	540
DB	481	CATGCTTTCAGCTTACGAGAGACTCGCGTATCGTCACTGCAATCCGACGTAGCGAA	540
QY	541	AGAGATTTCTGAATAGACCGGCTTTTGTGATGACCGGCTTAAAGATGAGCTTACTACT	600
DB	541	AGAGATTTCTGAATAGACCGGCTTTTGTGATGACCGGCTTAAAGATGAGCTTACTACT	600
QY	601	GATGTTTAAAGAGCAATTTGTTTGAACACAGGCTGTTTACTGCGCAAGCGTTGCGCG	660
DB	601	GATGTTTAAAGAGCAATTTGTTTGAACACAGGCTGTTTACTGCGCAAGCGTTGCGCG	660
QY	661	TATGCTTTCAGCAATCTCTTTTATGTAACAAATCAAGAGCGGCAAGCAAGCAAG	720
DB	661	TATGCTTTCAGCAATCTCTTTTATGTAACAAATCAAGAGCGGCAAGCAAGCAAG	720
QY	721	AGTATTCAGCAAGCAAGATGTTGATTTCTTGAAGAAAGAGATGATTAAGCAACCTGTTT	780
DB	721	AGTATTCAGCAAGCAAGATGTTGATTTCTTGAAGAAAGAGATGATTAAGCAACCTGTTT	780

QY 781 TGTTCGTGAGTTCCTTAAACGGCGCTCGCTTAAACAATGATGCTGTATTCGACA 840
DB 781 TGTTCGTGAGTTCCTTAAACGGCGCTCGCTTAAACAATGATGCTGTATTCGACA 840
QY 841 AGAGTATGAGCTTGAAAAAACAATGTGAGTATGAGTAAATGTGAGAAAGTTATTA 900
DB 841 AGAGTATGAGCTTGAAAAAACAATGTGAGTATGAGTAAATGTGAGAAAGTTATTA 900
QY 901 TTTGCTCGAAGCTTGATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 960
DB 901 TTTGCTCGAAGCTTGATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 960
QY 961 AAGACTCCGCTTGATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 1020
DB 961 AAGACTCCGCTTGATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 1020
QY 1021 TATATCGGACACCGTATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1080
DB 1021 TATATCGGACACCGTATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1080
QY 1081 CTCCCTCCATGATGATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 1140
DB 1081 CTCCCTCCATGATGATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 1140
QY 1141 CACACATTTATTTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1200
DB 1141 CACACATTTATTTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1200
QY 1201 TTAGATTAAACATGAAATTTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 1260
DB 1201 TTAGATTAAACATGAAATTTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 1260
QY 1261 AGAGCTTTTGTGATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 1320
DB 1261 AGAGCTTTTGTGATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 1320
QY 1321 AAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1380
DB 1321 AAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1380
QY 1381 AACAAGTGGGCTTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1440
DB 1381 AACAAGTGGGCTTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1440
QY 1441 AATCAAGCTGACAAACGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1500
DB 1441 AATCAAGCTGACAAACGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1500
QY 1501 CTGAGTGGCTTACCTTCAATATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1560
DB 1501 CTGAGTGGCTTACCTTCAATATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1560
QY 1561 CTCAGGCGCACTTCTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1620
DB 1561 CTCAGGCGCACTTCTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1620
QY 1621 GTCTTGTTCGGGACGAGGACCAAGATGAGTATGAGTATGAGTATGAGTATGAGTAT 1680
DB 1621 GTCTTGTTCGGGACGAGGACCAAGATGAGTATGAGTATGAGTATGAGTATGAGTAT 1680
QY 1681 ACGTGTGGGCTGATCTTGTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 1740
DB 1681 ACGTGTGGGCTGATCTTGTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 1740
QY 1741 TGAAGTTTTCGGTTCCTTGTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 1800
DB 1741 TGAAGTTTTCGGTTCCTTGTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 1800
QY 1801 TTTGCCCCGGGAGAGATCTTGTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 1860
DB 1801 TTTGCCCCGGGAGAGATCTTGTGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 1860
QY 1861 ATGAGTTTGAATGGGACCGCTTCAATGAGTATGAGTATGAGTATGAGTATGAGTATG 1920

DB 1861 ATGAGTTTGAATGGGACCGCTTCAATGAGTATGAGTATGAGTATGAGTATGAGTATG 1920
QY 1921 TTTCTTGAGAGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1980
DB 1921 TTTCTTGAGAGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGT 1980
QY 1981 AAGAAGCTCATATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 2009
DB 1981 AAGAAGCTCATATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 2009

RESULT 2
US-09-349-385-5
Sequence 5, Application US/09349385
Patent No. US20020152495A1
GENERAL INFORMATION:
APPLICANT: Ito, Toshio
APPLICANT: Fromm, Michael
APPLICANT: Meyerowitz, Elliot
TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT
FILE REFERENCE: MEI-0002
CURRENT APPLICATION NUMBER: US/09/349,385
CURRENT FILING DATE: 1999-07-09
EARLIER APPLICATION NUMBER: 60/115,967
EARLIER FILING DATE: 1999-01-15
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 5
LENGTH: 1593
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: cDNA
US-09-349-385-5

Query Match 48.6%; Score 976; DB 10; Length 1593;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 ATGCTAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 220
DB 1 ATGCTAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 60
QY 221 AGCCAAACCACTTCCCTTCTCCCTCTCCCTCTCCCTCTCCCTCTCCCTCTCCCTCTCT 280
DB 61 AGCCAAACCACTTCCCTTCTCCCTCTCCCTCTCCCTCTCCCTCTCCCTCTCCCTCTCT 120
QY 281 CTCTCTTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 340
DB 121 CTCTCTTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 180
QY 341 ATATCCGCTTCAATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 400
DB 181 ATATCCGCTTCAATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 240
QY 401 GGAAGATGATCTATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 460
DB 241 GGAAGATGATCTATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 300
QY 461 AAATTCGAGCCAAAGGCTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 520
DB 301 AAATTCGAGCCAAAGGCTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 360
QY 521 TCGAATCCGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 580
DB 361 TCGAATCCGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
QY 581 AAGAATGAGCTTATCTATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 640
DB 421 AAGAATGAGCTTATCTATGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 480
QY 641 TACTGCGAAGCGCTTCCCGATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATG 700

Db 481 TACTGCGAAGCGTTGCGCGTATCGTTGGAACCACTCTTAGTAAACAAATTCAGA 540
 Qy 701 AGAGCCGAGACCAACGACGAGTATCTCAACCCAGATGTTGATTTCTGAAAAAG 760
 Db 541 AAGGCGAGAGCGCAACGACGAGTATCTCAACCCAGATGTTGATTTCTGAAAAAG 600
 Qy 761 AGTAGTAAACGACCGCTGTTTGTGTGAGTGTCTTAAACCGCGCTGCTTAAACAATG 820
 Db 601 AGTAGTAAACGACCGCTGTTTGTGTGAGTGTCTTAAACCGCGCTGCTTAAACAATG 660
 Qy 821 ATGTGCTGTATTCGAGCAAGATATGAGTTGAGTTGAAAAACCATGTTGAGTTAGTAA 880
 Db 661 ATGTGCTGTATTCGAGCAAGATATGAGTTGAGTTGAAAAACCATGTTGAGTTAGTAA 720
 Qy 881 ATGTGCAAGAAAGTTATGATTTGCTCGGACGTTGATGATGATCACTTCTTGG 940
 Db 721 ATGTGCAAGAAAGTTATGATTTGCTCGGACGTTGATGATGATCACTTCTTGG 780
 Qy 941 CTATCGGAGTTGATCTCAAGACGTCGGCTAGATGTTCCACTGCTGACCAAGGA 1000
 Db 781 CTATCGGAGTTGATCTCAAGACGTCGGCTAGATGTTCCACTGCTGACCAAGGA 840
 Qy 1001 AACCGGTTTGTATCCCGGATTAATCCGAAACCGTAAATCAACCGGTGATTTGCTGCT 1060
 Db 841 AACCGGTTTGTATCCCGGATTAATCCGAAACCGTAAATCAACCGGTGATTTGCTGCT 900
 Qy 1061 GATTTCGAGCGTTTGTGCTCCCTCCATGCTGATGATTAATTCGACCCGAGCAATA 1120
 Db 901 GATTTCGAGCGTTTGTGCTCCCTCCATGCTGATGATTAATTCGACCCGAGCAATA 960
 Qy 1121 ATGCGCGTTCTTTGGG 1136
 Db 961 ATGCGCGTTCTTTGGG 976

RESULT 3

US-10-022-025A-2
 ; Sequence 2, Application US/10022025A
 ; Publication No. US20030092014A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Coleman, John R.
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Polypeptides for Catabolism of Abscisi-
 ; FILE OF INVENTION: Acid
 ; FILE REFERENCE: 3310 0003
 ; CURRENT APPLICATION NUMBER: US/10/022,025A
 ; CURRENT FILING DATE: 2003-01-09
 ; PRIOR APPLICATION NUMBER: US 60/254,819
 ; PRIOR FILING DATE: 2000-12-13
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: Patent version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 1593
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1593)
 US-10-022-025A-2

Query Match 48.6%; Score 976; DB 15; Length 1593;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 976; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 161 ATGCTTCGAAACCTCGAAAGCTCTTAACTTTCGCTTTTTCGCAATGCAAGCTTCTA 220
 Db 1 ATGCTTCGAAACCTCGAAAGCTCTTAACTTTCGCTTTTTCGCAATGCAAGCTTCTA 60
 Qy 221 AGCAAAACCAACCTTGGCTTCTCCCTCCTCGCGTCAATCAATGCTGCGCATATCT 280
 Db 61 AGCAAAACCAACCTTGGCTTCTCCCTCCTCGCGTCAATCAATGCTGCGCATATCT 120
 Qy 281 CTCCTTATGAACTATCCCGGTGAGCTGCTTGGGGAAATACCTCTTCCGCGGTTA 340

Db 121 CTCCTTATGAACTATCCCGGTGAGCTGCTTGGGGAAATACCTCTTCCGCGGTTA 180
 Qy 341 ATATCCGCTTCATACAAAACCGGAAAGCTTATTCGCGTCCAAAAGGCTTCCCTTGGTT 400
 Db 181 ATATCCGCTTCATACAAAACCGGAAAGCTTATTCGCGTCCAAAAGGCTTCCCTTGGTT 240
 Qy 401 GGAAGCATGTCATCATGTCAGACACTAGCTCAGCGACGATTCGTCAGCTGAG 460
 Db 241 GGAAGCATGTCATCATGTCAGACACTAGCTCAGCGACGATTCGTCAGCTGAG 300
 Qy 461 AAATTCGAGCCAAAGAGGCTCATGCTTCAAGCTTGAAGAGACTGCGGTATGTCAG 520
 Db 301 AAATTCGAGCCAAAGAGGCTCATGCTTCAAGCTTGAAGAGACTGCGGTATGTCAG 360
 Qy 521 TGCAATCCGACGTAAGGAAAGATTCGAAATAGCCCGGTTTGTGATGACCGGTT 580
 Db 361 TGCAATCCGACGTAAGGAAAGATTCGAAATAGCCCGGTTTGTGATGACCGGTT 420
 Qy 581 AAAGATCGGCTTACTCACTGATGTTTAAACAGACAAATGTTTGAACCAACGAGTT 640
 Db 421 AAAGATCGGCTTACTCACTGATGTTTAAACAGACAAATGTTTGAACCAACGAGTT 480
 Qy 641 TACTGCGAAACGCTTCCGCTATGCTTGCAGACACTCTCTTATGTAACAAACAAATCAGA 700
 Db 481 TACTGCGAAACGCTTCCGCTATGCTTGCAGACACTCTCTTATGTAACAAACAAATCAGA 540
 Qy 701 AGAGCCGAGACGCAACGACGAGTATCTCAAGCCAGATGAGTTTCTTGAAGAAAAAG 760
 Db 541 AGAGCCGAGACGCAACGACGAGTATCTCAAGCCAGATGAGTTTCTTGAAGAAAAAG 600
 Qy 761 AGTAGTAAACGACCGCTGTTTGTGCTGAGTTGCTTAAACCGGCTGCTTAAACAATG 820
 Db 601 AGTAGTAAACGACCGCTGTTTGTGCTGAGTTGCTTAAACCGGCTGCTTAAACAATG 660
 Qy 821 ATGTGCTGTATTCGAGCAAGATATGAGTTGAGTTGAAAAACCATGTTGAGTTAGTAA 880
 Db 661 ATGTGCTGTATTCGAGCAAGATATGAGTTGAGTTGAAAAACCATGTTGAGTTAGTAA 720
 Qy 881 ATGTGCAAGAAAGTTATGATTTGCTCGGACGTTGATGATTAATTCGACCCGAGCAATA 940
 Db 721 ATGTGCAAGAAAGTTATGATTTGCTCGGACGTTGATGATTAATTCGACCCGAGCAATA 780
 Qy 941 CTATCGGAGTTGATCTCAAGACGTCGGCTAGATGTTCCACTGCTGACCAAGGA 1000
 Db 781 CTATCGGAGTTGATCTCAAGACGTCGGCTAGATGTTCCACTGCTGACCAAGGA 840
 Qy 1001 AACCGGTTTGTATCCCGGATTAATCCGAAACCGTAAATCAACCGGTGATTTGCTGCT 1060
 Db 841 AACCGGTTTGTATCCCGGATTAATCCGAAACCGTAAATCAACCGGTGATTTGCTGCT 900
 Qy 1061 GATTTCGAGCGTTTGTGCTCCCTCCATGCTGATGATTAATTCGACCCGAGCAATA 1120
 Db 901 GATTTCGAGCGTTTGTGCTCCCTCCATGCTGATGATTAATTCGACCCGAGCAATA 960
 Qy 1121 ATGCGCGTTCTTTGGG 1136
 Db 961 ATGCGCGTTCTTTGGG 976

Search completed: January 16, 2004, 21:27:40
 Job time : 663 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 00:13:09 ; Search time 450 Seconds
(without alignments)
9536.011 Million cell updates/sec

Title: US-10-022-025a-2

Perfect score: 1593

Sequence: 1 atggtcagcgaactcgaag.....tcgcgcgtagcgscagctaa 1593

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	976	61.3	1902 25	ABX10941
2	975.6	61.2	2005 21	AACT9902
3	794	49.8	25	ABX10942
4	651.6	40.9	20	AAK60776
5	599.6	37.6	25	ABT23239
6	585.8	36.8	25	ABT23241
7	355	22.3	20	AAK55600
8	352	22.1	25	ABT23246

9	316.6	19.9	1934 25	ABT23240	Seed development e
10	290.6	18.2	1687 25	ABT23249	Seed development e
11	286	18.0	1545 25	ABT23284	Seed development e
12	283	17.8	1758 25	ABT23283	Seed development e
13	278.2	17.5	1597 25	ABT23237	Seed development e
14	275.2	17.3	1585 25	ABT23282	Seed development e
15	274.4	17.2	1539 25	ABT23238	Seed development e
16	265.8	16.7	2924 25	ABT23242	Seed development e
17	234	14.7	1146 25	ABT23235	Seed development e
18	233.4	14.7	1578 25	ABT23228	Seed development e
19	213.4	13.4	1131 25	ABT23233	Seed development e
20	194.8	12.2	1557 25	ABT23285	Seed development e
21	189.2	11.9	884 25	ABT23247	Seed development e
22	156.4	9.8	442 25	ABT23248	Seed development e
23	155.6	9.8	1701 25	ABT23229	Seed development e
24	144.2	9.1	1704 20	AAK60781	Soybean cytochrome
25	141.8	8.9	1789 18	AAK60655	Pectunia flavonoid
26	135.6	8.5	1667 18	AAK94664	Gentian flavonoid
27	129.6	8.1	1708 20	AAK81873	DNA encoding a P45
28	129.6	8.1	1708 21	AAK86881	Sweetgum cytochrom
29	129.6	8.1	1708 22	AAK11125	L. styraciflua ang
30	126	7.9	778 25	ABT23236	Seed development e
31	124.4	7.8	1660 18	AAK94661	Chrysanthemum flav
32	124.4	7.8	1748 18	AAK94660	Rose flavonoid 3'-
33	121.2	7.6	1745 18	AAK94656	Carnation flavonoi
34	119.4	7.5	1527 24	ABK12518	DNA encoding p-con
35	119.4	7.5	1549 24	ABK12520	DNA construct for
36	117.8	7.4	1527 24	ABK12519	DNA encoding modif
37	117.4	7.4	1711 18	AAK94657	Snaptagon flavono
38	113.8	7.1	1214 18	AAK94655	Lisiantus flavono
39	111.8	7.0	1121 21	AAK36642	Arabidopsis thalia
40	111.8	7.0	1815 18	AAK94662	Toronia flavonoid
41	109.2	6.9	1824 18	AAK94663	Morning glory flav
42	106.8	6.7	1770 22	AAK30281	Perilla flavone sy
43	106.8	6.7	1933 25	ABK93222	cDNA encoding cot
44	105.6	6.6	457 25	ABT23244	Seed development e
45	104	6.5	618 22	AAK87696	Peppermint plant o

ALIGNMENTS

RESULT 1	ABX10941	Arabidopsis thaliana
ID	ABX10941 standard; cDNA, 1902 BP.	
XX		
AC	ABX10941;	
XX		
DT	16-APR-2003 (first entry)	
XX		
DE	cDNA encoding Arabidopsis thaliana cytochrome P450 protein.	
XX		
KW	Arabidopsis; cytochrome P450; CP450; plant; parthenocarpic; fruit;	
KW	enlarged fruit; transgenic; vegetable size; grain size; gene; ss;	
KW	leaf size; flower size; agricultural yield; male-sterile plant.	
XX		
OS	Arabidopsis thaliana.	
XX		
FN	Key	Location/Qualifiers
FT	CDS	69..1673
FT		/tag= a
FT		/product= "A. thaliana cytochrome P450"
XX		
PN	US2002152495-A1.	
XX		
PD	17-OCT-2002.	
XX		
PF	09-JUL-1999;	99US-0349385.
XX		
PR	15-JAN-1999;	99US-115967P.
XX		
PA	(ITOP/) ITO T.	
PA	(FROM/) FROM M.	

PA (MEYE/) MEYEROWITZ E.
 XX Ito T, Fromm M, Meyerowitz E;
 XX WPI, 2003-198390/19.
 DR P-PSDB; AB575612.

XX New polynucleotides encoding a cytochrome P450 polypeptide, useful for
 PT producing plants with enlarged or parthenocarpic fruits, in increasing
 PT plant tissue size to increase agricultural yields, or producing
 PT male-sterile plants

XX Claim 4, Page 12-13, 31pp, English.

CC This invention relates to an isolated polynucleotide encoding a
 CC cytochrome P450 polypeptide which when expressed in a plant produces at
 CC least one phenotype selected from parthenocarpic fruit and enlarged
 CC fruit. The invention also discloses a recombinant construct comprising
 CC the novel polynucleotide which when expressed in a plant produces a
 CC plant with a parthenocarpic fruit or enlarged fruit. Also disclosed is
 CC a transgenic plant which expresses a cytochrome P450 or comprising
 CC an isolated polynucleotide encoding a cytochrome P450 polypeptide which
 CC when expressed in a plant produces a plant with a parthenocarpic fruit
 CC or enlarged fruit compared with a plant lacking the isolated
 CC polynucleotide. The cytochrome P450 polypeptide and polynucleotide of
 CC the invention is useful in the production of plants with seedless,
 CC enlarged, or parthenocarpic fruits, including vegetable or grain size,
 CC leaf size or flower size. The sequences may also be used in increasing
 CC plant tissue size to increase agricultural yields of plants, to produce
 CC male-sterile plants and to screen for compounds that control
 CC parthenocarpic or fruit size in plants. The present sequence represents
 CC the cDNA sequence encoding the Arabidopsis thaliana cytochrome P450
 CC protein of the invention.

XX Sequence 1902 BP, 507 A; 448 C; 439 G; 508 T; 0 other;

XX Query Match 61.3%; Score 976; DB 25; Length 1902;
 XX Best Local Similarity 78.0%; Pred. No. 0;
 XX Matches 128; Conservative 0; Mismatches 320; Indels 30; Gaps 4;

QY 19 AGCTCTTATCTTTTCCCTTTTGTCCAAATGACGGCTCTAAGCCAAACCACTTGGC 78
 DB 90 AGTAGCTTACTTTTGGCCCTCTGTGTCAAATGTAGCTCTTACTCAACCAATCTTCT 149
 QY 79 TTCTCCCTCGCCGCTGCAATCATCTGCTGCGCAATCTCTCTTATGACCTAT 138
 DB 150 CTCTCTCTCTGTAACCTCCAGCTTCTCTGCTCTTCTCTCTCTCTGCTCAT 209
 QY 139 CCCGCTGACCTGCTGGGGAATACTCTTGGCGCGGTATATCCGCTTCATCAAA 198
 DB 210 CCCGAGGACCCCGCATGGGGAATACTCTCTCCACCGC-----CGCGCT 254
 QY 199 ACCGAAAGCTTATCCCGGTCCAAAGGCTTCCCTTGTGGAAGATGCACTCATG 258
 DB 255 CAACACACCTGATATCCCGGCAAGAGCTTACTTTTGTGGAAGATGCTCTCATG 314
 QY 259 TCAAGCACTCTAGCTCAACGAGATGCTGATGACGTTGAGAAATTCGAGCCAAAGG 318
 DB 315 TCAAAACATCTGCTCAACGTTGCATGACGCAACGAGAAATTTAGACCCGAAACG 374
 QY 319 CTGATGCTTCAAGCTTGAAGAGACTCGGCTGATGTCACGCGCAATCCGACGATGAG 378
 DB 375 TTATGAGCTTTAGTTTGGAGAACTCGGCTGATGTCACGCGCAATCTGATGAGCT 434
 QY 379 AAAGAAATCTGAATAGCCCGTTTGTCTGATGCAACGCTTAAAGATCGCTTATCTCA 438
 DB 435 AAAGAAATCTGAATAGCCCGTTTGTCTGATGCAACGCTTAAAGATCGCTTATCTCC 494
 QY 439 CTGATGCTTCAAGCTTGAAGAGACTCGGCTGATGTCACGCGCTTAAAGATCGCTTATCT 498
 DB 495 CTGATGCTTCAAGCTTGAAGAGACTCGGCTGATGTCACGCGCTTAAAGATCGCTTATCT 554
 QY 499 CGTATGCTTCAAGCTTGAAGAGACTCGGCTGATGTCACGCGCTTAAAGATCGCTTATCT 558

DB 555 AAATGCGCTTAAATCATCTTTTACGCCGGAACGATTTAAACGTTCCGAACGCGAGGA 614
 QY 559 CGAGTATCTCAAGCAGATGTTGATTTCTTGAAGAAACAGATGATGACGA---CCG 615
 DB 615 AGCGTATCGCAATCAATCGTGAAGTGTCTCACAAAACAGATGACCAAGGCTCTC 674
 QY 616 TGTTTTGTCTGAGTTGCTTAAACGGCGTGTCTTAAACATGATGCTCTGATTC 675
 DB 675 TGTTCGCAAGTACTGATCAAAACGGCATGCTTAAATGATGATGCTCTGATTC 734
 QY 676 GGAAGAGATGAGCTTGAAGAAACATG-----TTAGTTAGTGAATGCTC 726
 DB 735 GGAAGAAATTCAGCTTGAAGAGATGAGAAAGTGAAGTGAAGTGAATGCTC 794
 QY 727 GAAGAAAGTATGATTTGCTCGAAGCTTGAATGATGATGATGATGATGATGATG 786
 DB 795 GAAGAAAGTATGATTTGCTCGAAGCTTGAATGATGATGATGATGATGATGATG 854
 QY 787 GAATTTGATCTTAAAGATCTCGGCTGATGATGATGATGATGATGATGATGATG 846
 DB 855 GAATTTGATCTTAAAGATCTCGGCTGATGATGATGATGATGATGATGATGATG 914
 QY 847 TTTGTATCCCGATTTATTCGAAACACCGTATCAACCGGATTTGCTGATGATTC 906
 DB 915 TTTGTATCCCGATTTATTCGAAACACCGTATCAACCGGATTTGCTGATGATTC 974
 QY 907 GTGAGCTTTTGTCTTCCCTCCATGTTGATGATGATGATGATGATGATGATGATG 966
 DB 975 GTTGAAGTATGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1034
 QY 967 GTTCTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1026
 DB 1035 GTTCTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1094
 QY 1027 CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1086
 DB 1095 CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1154
 QY 1087 GTAGTGGGAATTCAGAGAGCTTGAATGATGATGATGATGATGATGATGATGATG 1146
 DB 1155 ATGCTGGAAGATCAAGAGCTGTCGAAGTGTGATGATGATGATGATGATGATGATG 1214
 QY 1147 GCTGTGTGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1206
 DB 1215 GCTGTGTGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1274
 QY 1207 TTGGCCATTAACAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1266
 DB 1275 TTGAATTCACAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1334
 QY 1267 GTGAACATGTTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1326
 DB 1335 GTGAACATGTTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1394
 QY 1327 CTTGAAGCTTGTGGAAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1386
 DB 1395 CTTGAAGCTTGTGGAAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1454
 QY 1387 AGACTTCACTTGTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1446
 DB 1455 AGCTTCACTTGTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1514
 QY 1447 ACCGTTATGTTTGAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1503
 DB 1515 ACCGTTATGTTTGAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1574
 QY 1504 GTTAACGCGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1563
 DB 1575 GAGAAGACCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1634
 QY 1564 CTTGCTAATTTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1591

```
DB      1635 GCTGCTAATACGCCCGACGCGCACTT 1662

RESULT 2
AAC47902
ID      AAC47902 standard; DNA; 2005 BP.
AC      AAC47902;
XX
XX      18-OCT-2000 (first entry)
DT
XX
XX      Arabidopsis thaliana DNA fragment SEQ ID NO: 55538.
DE
XX
XX      Hybridisation assay; genetic mapping; gene expression control;
KM      protein identification; signal transduction pathway;
KW      metabolic pathway; promoter; termination sequence; ss.
XX
XX      Arabidopsis thaliana.
OS
XX
XX      EP1033405-A2.
PN
XX
XX      06-SEP-2000.
PD
XX
XX      25-FEB-2000; 2000EP-0301439.
PF
XX
XX      25-FEB-1999; 99US-0121825.
XX      05-MAR-1999; 99US-0123180.
PR      09-MAR-1999; 99US-0123548.
PR      23-MAR-1999; 99US-0125788.
PR      25-MAR-1999; 99US-0126264.
PR      29-MAR-1999; 99US-0126785.
PR      01-APR-1999; 99US-0127462.
PR      06-APR-1999; 99US-0128224.
PR      08-APR-1999; 99US-0128714.
PR      16-APR-1999; 99US-0129845.
PR      19-APR-1999; 99US-0130077.
PR      21-APR-1999; 99US-0130449.
PR      23-APR-1999; 99US-0130510.
PR      28-APR-1999; 99US-0130891.
PR      30-APR-1999; 99US-0131449.
PR      30-APR-1999; 99US-0132048.
PR      30-APR-1999; 99US-0132407.
PR      04-MAY-1999; 99US-0132484.
PR      05-MAY-1999; 99US-0132485.
PR      06-MAY-1999; 99US-0132486.
PR      06-MAY-1999; 99US-0132487.
PR      07-MAY-1999; 99US-0132863.
PR      11-MAY-1999; 99US-0134256.
PR      14-MAY-1999; 99US-0134218.
PR      14-MAY-1999; 99US-0134219.
PR      14-MAY-1999; 99US-0134221.
PR      14-MAY-1999; 99US-0134370.
PR      18-MAY-1999; 99US-0134768.
PR      19-MAY-1999; 99US-0134941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135353.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139493.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139753.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      01-JUL-1999; 99US-0142154.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      19-JUL-1999; 99US-0144335.
PR      19-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144864.
PR      20-JUL-1999; 99US-0144864.
PR      21-JUL-1999; 99US-0144864.
PR      21-JUL-1999; 99US-0145086.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145218.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
PR      27-JUL-1999; 99US-0145919.
PR      28-JUL-1999; 99US-0145951.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
PR      02-AUG-1999; 99US-0146389.
PR      03-AUG-1999; 99US-0147038.
PR      04-AUG-1999; 99US-0147204.
PR      04-AUG-1999; 99US-0147302.
PR      05-AUG-1999; 99US-0147192.
PR      05-AUG-1999; 99US-0147260.
PR      06-AUG-1999; 99US-0147303.
PR      06-AUG-1999; 99US-0147416.
PR      09-AUG-1999; 99US-0147493.
PR      09-AUG-1999; 99US-0147935.
PR      10-AUG-1999; 99US-0148171.
PR      11-AUG-1999; 99US-0148319.
PR      12-AUG-1999; 99US-0148341.
PR      13-AUG-1999; 99US-0148565.
PR      13-AUG-1999; 99US-0148684.
PR      16-AUG-1999; 99US-0149368.
PR      17-AUG-1999; 99US-0149375.
```

PR	18-AUG-1999	99US-0149426
PR	20-AUG-1999	99US-0149722
PR	20-AUG-1999	99US-0149723
PR	20-AUG-1999	99US-0149929
PR	23-AUG-1999	99US-0149902
PR	23-AUG-1999	99US-0149930
PR	25-AUG-1999	99US-0150566
PR	26-AUG-1999	99US-0150884
PR	27-AUG-1999	99US-0151065
PR	27-AUG-1999	99US-0151066
PR	27-AUG-1999	99US-0151080
PR	30-AUG-1999	99US-0151303
PR	31-AUG-1999	99US-0151438
PR	01-SEP-1999	99US-0151930
PR	07-SEP-1999	99US-0152263
PR	10-SEP-1999	99US-0153070
PR	13-SEP-1999	99US-0153758
PR	15-SEP-1999	99US-0154018
PR	16-SEP-1999	99US-0154039
PR	20-SEP-1999	99US-0154779
PR	22-SEP-1999	99US-0155139
PR	23-SEP-1999	99US-0155386
PR	24-SEP-1999	99US-0155539
PR	28-SEP-1999	99US-0156458
PR	29-SEP-1999	99US-0156596
PR	04-OCT-1999	99US-0157717
PR	05-OCT-1999	99US-0157753
PR	06-OCT-1999	99US-0157865
PR	07-OCT-1999	99US-0158029
PR	08-OCT-1999	99US-0158322
PR	12-OCT-1999	99US-0158363
PR	13-OCT-1999	99US-0159293
PR	13-OCT-1999	99US-0159294
PR	13-OCT-1999	99US-0159295
PR	14-OCT-1999	99US-0159329
PR	14-OCT-1999	99US-0159330
PR	14-OCT-1999	99US-0159331
PR	14-OCT-1999	99US-0159637
PR	14-OCT-1999	99US-0159638
PR	18-OCT-1999	99US-0159584
PR	21-OCT-1999	99US-0160741
PR	21-OCT-1999	99US-0160767
PR	21-OCT-1999	99US-0160768
PR	21-OCT-1999	99US-0160770
PR	21-OCT-1999	99US-0160814
PR	21-OCT-1999	99US-0160815
PR	22-OCT-1999	99US-0160980
PR	22-OCT-1999	99US-0160981
PR	22-OCT-1999	99US-0160989
PR	25-OCT-1999	99US-0161404
PR	25-OCT-1999	99US-0161405
PR	25-OCT-1999	99US-0161406
PR	26-OCT-1999	99US-0161359
PR	26-OCT-1999	99US-0161360
PR	26-OCT-1999	99US-0161361
PR	28-OCT-1999	99US-0161920
PR	28-OCT-1999	99US-0161992
PR	28-OCT-1999	99US-0161993
PR	29-OCT-1999	99US-0162142

	Query Match	61.2%;	Score 975.6;	DB 21;	Length 2005;
	Best Local Similarity	77.9%;	Ped. 0.0;		
	Matches 123/;	Conservative	1;	Mismatches 320;	Indels 30; Gaps 4
Qy	19	AGCTCTTAATCTTTGCCCTTTTGTGCCAATAGCAGCGCTTCTTAGGCCAAACCACTTGGC	78		
Db	199	AGTAGCTTACTTTTGGCCCTCTTGTGCCAATATAGCCCTCTTATCTCAACCAACTTGGCT	258		
Qy	79	TTCTCCCTCTCTGCGCGGTCAACATCATCTGGCTGGCCATATCTCTTTATATGACCTAT	138		
Db	259	CTCTCTCTCTCTGAGCTCCCTAGCTTCTCTGGCTCTTCTCTCTTCTTCTGGCTCAT	318		

Dp	319	CCGGAGAGACCGGATGGGAAAATACCTCTCTCCACCGC-----GGCCGT	363
Qy	199	ACCGGAAAGCTTATTCGCCGTCCAAAGGCTTCCCTTTGGTTGGAAGCATGTCACTCATG	258
Dp	364	CAACACACCGTATATCCCGGGCAAGAGGCTTACCTTTTGTCCGAAGCATGTCTTCATG	423
Qy	259	TCAGCACTCTAGCTCACCGGACGAATCGCTGATGCAAGCTGAGAAATTCCGAGCCAAAGAG	318
Dp	424	TCAAACACTCTGGCTCACCGTTGCATAGCCGAAACCGCAGAGAAATTTAGACCCGAA	483
Qy	319	CTCATGCTTTTACGTTTAGGAGAGACATCGCGGTATGTCACGTCGCAATCCCGAGCTTAGC	378
Dp	484	TTATGCGCGTTAGTTTGGAGAAATCGCGGTATGTCACGTCGCAATCTGATGTACT	543
Qy	379	AAAGAGATTTCTGAATAGCCCGTTTTTGTCTGATCGACCGGTTTAAAGATCGGCTTACTCA	438
Dp	544	AAAGAGATTTCTAAACAGTCCGGTTTTGCTGACGCGCCGGTTAAGGAATCAGCTTATTC	603
Qy	439	CTGATGTTTAAACAGAGCAATGTGTTTTGACACACAGGTGTTTACGCGGAACCTTGAC	498
Dp	604	CTCATGTTTAAACCGTGCATTCGGTTTTGCTCTTACGGCGTTTACTCGGGAACCTTAGGA	663
Qy	499	CGTATGCTTGCACATCTCTTTAGTACAAAACAATCAGAGAGCGGACGACACGA	558
Dp	664	AAATGCGCTTAAATCACTTTTACGCGCAAAACAATTTAAAGTTCCGAAACGACAGGA	723
Qy	559	CGAGTATCTCAGCGCAATGGTTGATTTCTTGAAGAAAACAGATGATACGA---CCG	615
Dp	724	AGGTATCGGAAATCAATCGTAGAGTCTCAAAAACAGATTAACCAAAAGTCTC	783
Qy	616	TGTTTTGTTCCGATATTCCTTAAACGGCGCGCTTAAACAATATGTGCTCTGATTC	675
Dp	784	TGTTTGACGAGCTTATCAAAACGGCAGCTTAAATCAATGATGTGCTCTGTGTTTC	843
Qy	676	GGAACAAGTATAGCTTGAAAAAAACCATG-----TTGATTAAGTAAATGSGTC	726
Dp	844	GGAAAAAGATACGAGCTTGAAGAAACAGAAAGATGAGAGTACGTAAATTGSGTG	903
Qy	727	GAAAGAGTATAGTTTCTCGAAGCTGAAATTTGACGTGATCACTTCTTGCGTATGCG	786
Dp	904	GAAAGAGTATAGTTTCTCGTACACTGAATTTGACCGATCATCTCCATGSGTCTCT	963
Qy	787	GAGTTGATCTCAAAACACTCCGGCTTAGATGTTTCACACTGTATCAAAAGGTTAAACGG	846
Dp	964	GAAATTGATCTCAAAAGATCCGGCTTAGATGTTTCACTGCTCCCAAAAGTTAAACCGG	1023
Qy	847	TTTGATCCCGGATTTATTCGGAACACCGTATCAAAACGGGTATTTGCTCGTATTC	906
Dp	1024	TTTGTAACCGGATTTATCTTGACCAACCGTAAACAACTCGTACACCGAGTACTTC	1083
Qy	907	GTGCAAGTTTTCCTCCCTCCGACAGTGTCAATTAATTTCCGACCCCGGACATAATGCGC	966
Dp	1084	GTGACAGTATGCTCTCTCTGATGAGTCTGATTAATTTTCCGACCTGATATCATGCGC	1143
Qy	967	GTTCTTTGGAGATGATTTCAAGAGAAACACACATGGTCCGCTTTATTCGAGTGATC	1026
Dp	1144	GTTCTTATGGGAATGATTTTCAAGAGAACTGACACCGTGGCTGTTTTGATTCAGGTGATTT	1203
Qy	1027	CTGCGTAGGATGGTCTTCATCGAGATATCAATCAACGGTATCAAAAACAGCTGATCAAA	1086
Dp	1204	CTTGCTGAGATGCTCTTCACTCCAGATTTCAATTCACGCTTCACAATGAGCTTGATCAA	1263
Qy	1087	GTAATGCGGAATCAAGAGCCCTTAGATGAATCTGACTTGCTTCACTTCCATATTTAACG	1146
Dp	1264	ATGCTGGGACGATCAAGGGCTGTGAAAGATCTGAAGTGGTGTCTCTGATATATTTAACG	1323
Qy	1147	GCTGTGTAAAGAAATTTGAGGCTTCATCTCCACAGGCGCACTTCTATCATGAGGCCGT	1206
Dp	1324	GCTGTGTAAAGAAATCTTAGGCTTCAACCGCACAGGCCCACTACTCTCATGAGGCCGT	1383
Qy	1207	TTGSCATTAACGACAGATGCTGATGAGTCTGCTTTGTTCCGACGGGACCAACAGCAATG	1286

Db	1384	TTAGCAATCACAGACAGATCATCGAGGTCGTGTCGCGCGG99AACACCGCAATG	1443
QY	1267	GTGAACATGTGGGCGGTATCGCATGATCCACACGTGTGGGTTGATCCTTTGGAGTTTAA	1326
Db	1444	GTGAACATGTGGGCTTATGTGACACAGATCCACACTGTGGGGAATCCGTGGAGTTTAA	1503
QY	1327	CCTGAGAGTTCGTGGCAAAAGAAGTGAAGTTTGATCTTGAGTTCGTGGTCCGATTTG	1386
Db	1504	CCCCACGTTTGTAGCCAGAGAGTGAAGTTGATCTTGAGTTCGTGGTCCGATTTG	1563
QY	1387	AGACTTGACACTTTCGGGTCCGGTCGTTCGATTTTGCCCGGAGAAATCTTGTTTACT	1446
Db	1564	AGGCTTGACCCGTTCCGGTCCGGTCGTTCGAGTTTGCCCGGAGAAATCTTGTTTGACC	1623
QY	1447	ACCGTTATGTTTGGACCGCGCATGATGTTACATGAATTTGAATG---GGACCGTCCGAT	1503
Db	1624	ACCGTTACGTTTGGACCTGCACCGCTTTTGATGATGTTGATGACTGCACCCGTCGAT	1683
QY	1504	GGTAAACGCGTTGACTTATCTGAGAACTGAAGCGTTTCTTCGAGATGGCTAATCCGTT	1563
Db	1684	GAGAAACCGGTGACTGTCCGAGAACTGAGCGTCTCTGTGAGATGGCTAATCCGTT	1743
QY	1564	CCTGCTAATATGCGCGCTAGCGCGCAATT	1591
Db	1744	GCTGCTAATTAACGCCCAAGCGCGCAATT	1771

	RESULT 3
ABX10942	
ID	ABX10942 standard; DNA; 7869 BP.
XX	
AC	ABX10942;
XX	
DT	16-APR-2003 (first entry)
XX	
DE	Arabidopsis thaliana cytochrome P450 gene sequence.
XX	
KW	Arabidopsis; cytochrome P450; CYP450; plant; parthenocarpic; fruit;
KW	enlarged fruit; transgenic; vegetable size; grain size; gene; ds;
KW	leaf size; flower size; agricultural yield; male-sterile plant.
XX	
OS	Arabidopsis thaliana.
XX	
PN	US2002152495-A1.
XX	
PD	17-OCT-2002.
XX	
PF	09-JUL-1999; 99US-0349385.
XX	
PR	15-JAN-1999; 99US-115567P.
XX	
PA	(ITOT/) ITO T.
PA	(FROM/) FROMM M.
PA	(MEYE/) MEYEROWITZ E.
XX	
PI	ITO T, FROMM M, Meyerowitz E;
XX	
DR	WPI; 2003-198390/19.

Claim 4; Page 13-16; 31pp; English

CC This invention relates to an isolated polynucleotide encoding a
CC cytochrome P450 polypeptide which when expressed in a plant produces at
CC least one phenotype selected from parthenocarpic fruit and enlarged
CC fruit. The invention also discloses a recombinant construct comprising
CC the novel polynucleotide which when expressed in a plant produces a
CC plant with a parthenocarpic fruit or enlarged fruit. Also disclosed is
CC a transgenic plant which expresses a cytochrome P450 or comprising

Sequence 7669 BP; 2628 A; 1415 C; 1374 G; 2452 T; 0 other;

Query Match	49.8%;	Score 794;	DB 25;	Length 7869;
Best Local Similarity	70.3%;	Pred. No. 9.6e-259;		
Matches 1238; Conservative	0;	Mismatches 320;	Indels 202;	Gaps 5;

QY 19 AGCTCTTAATCTTTGGCCCTTTTGTCGAAATGAGAGGTTCTAAGCCAAACCAACTCTGGC 78
 Db 2154 AGTAGCTTAATCTTTGGCCCTTTTGTCGAAATGAGAGGTTCTAAGCCAAACCAACTCTGGC 2213
 QY 79 TTCTCCCTCTCGCCGTCACAAATCATCTGGCTCGCCATATCTCTCTTTTAGACCTAT 138
 Db 2214 CTCTCTCTCTCGTAGCTCCCTAGGCTCTCTCGCTCTTCTCTCTCTCTCTGTGCTCAT 2273
 QY 139 CCGGTGAGACTGTGTTGGGGGAAATACCTCTTGGCGCGTTAATATCCGGTTCAATCAAA 198
 Db 2274 CCCGAGAGACCCCGATGAGAAAGTATTCTTCACCG-----CCGCGCT 2318
 QY 199 ACCGAAACGTTATTCGCCGTCCAAAAGCTTCCTCTTTGGTTGAAAGCATGTCACTCATG 258
 Db 2319 CAATCAACCGTGATACCCGSGCCAAAGAGGCTTACCTTTGTCCGAAGACATGCTCATG 2378
 QY 259 TCAAGCACTTAGTCTACCCGACGATGCTGTATGACGTGAGAAATTCGGAGCCAAAGG 318
 Db 2379 TCAAAACCTCGCTCACCGTTGATAGCCGCAACCGCAAGAAATTAAAGCCGAACGG 2438
 QY 319 CTCAATGCTTACAGTTAGAGAGACATCGCGTGATCGTCAAGTGCATCCCGAGTACG 378
 Db 2439 TTATGCGCTTATGTTTGGAGAACTCGGTATGCTCAAGTGCATCCGATGATGAGT 2498
 QY 379 AAAGAGATTCGAATAGCCCGGTTTTTGTGATGACCCGTTAAAGATGGGCTTACTCA 438
 Db 2499 AAAGAGATTCGAACAGTCCCGTTTTTGTGATGACCCCGGTTAAGAAATCAGCTTATTC 2558
 QY 439 CTGATGTTTACAGAGCAATGTTGTTTGACCAACAGGTGTTTACTAGCGAACGCTTCG 498
 Db 2559 CTCAATGTTAACCGTGCTCATCGGTTTTGCTCTTACCGCGTTTACGCGAACCTTGAGA 2618
 QY 499 CGTATCGCTTGAAACCATCTCTTTTAGTACAAACCAATCAGAGAGCCGAGACCGAACGA 558
 Db 2619 AAAATCGGCTTAATCATCTTTTACGCCCGCAACAGATTAAACCTTCCGAACCCAGAGA 2678
 QY 559 CGAGTGATCTCAAGCCAAATGTTGAGTTTCTTGAATAACGAGATGATTAACGA---CCG 615
 Db 2679 AGCGTGATCGGAATCAATCGTAGTAGTGTCTCACAAACAGATTAACCAAAAGTCTC 2738
 QY 616 TGTTTGTTGCGAGTTGCTTAAACCGCGCTCGCTTAAACAATGATGTGCTCTGTATTC 675
 Db 2739 TCTTTCGACGTAATTGATCAAAACGCGCATCGCTTATATACATGATGTGCTGTCTTTC 2798
 QY 676 GGAACAAGATGATGAGCTTGAATAAAACATG-----TTGAGTTAACGGAATGGTTC 726
 Db 2799 GGAATAAGATACGAGCTTGAGAGAAAGATAGAGATGAGAGCTACGTAATTTGATG 2855
 QY 727 GAAGAAGGTATGATTTGCTCGGAACGTTGATTTGAATGACTGATCACTTCTTGCTATCG 786
 Db 2859 GAAGAAGGTATGATTTGCTCGTACACTGATTTGACCGATCATCTCCCATGGCTCTCT 2918
 QY 787 GAGTTGATCTCAAAAGACTCGGCTCTAGATTTTCCACATCGTACCAAGATTAACCGG 846

QY 481 TACTGGCGAA CGCTTCGCCGATCGCTTCGAAACCATCTCTTAGTACAAAACAATCAGA 540
 DB 473 TACTGGCGAAAGCTCAGAGAAATCGCTTAATCACTTTCGCCCGCCGAGATATAA 532
 QY 541 AGAGCCGAGAGCAAGCAAGATGATCTCAAGCCAGATGGTGAAGTTCTTGAANAACAG 600
 DB 533 GCCTTAGCTCCCAAGCTCTCAAAATGCGGCCCAATGGGTACATCTT---AAATTAAC 589
 QY 601 AGTAGTAACGAAACCTGTTTGTTCGTGAGTTGCTTAAACGGCGTTCGCTTAAACAATG 660
 DB 590 AAGGCCACCGCAGCTTACGTTGTCGCAAGTGTGAAAAGGCTTCGCTCACTAATCAG 649
 QY 661 ATGTGCTCTGATTCGCAACAAGATAGCTTGAAAAAACA-----TGTAG 711
 DB 650 ATGTGCTCCGTTTGGACAGAGTATAGCTGACGACCCAAACAGCCGAAATGAAAGAC 709
 QY 712 TTACGTAAATGTCGAAAGAGTTATGATTTGCTGGAACGTTGAATTGACTGATAC 771
 DB 710 CTGGAAATATTAGTACCAAGTTATGACCTGTGGGCTGTATTATGGGCGACAC 769
 QY 772 CTTCCTTGCTATCGAGTTGATCTCAAAAGCTCCGCTTGAAGTTCCACACTGTA 831
 DB 770 CTTCCTTTCTTGACATTCGACGCCCAAAATATCGGTTCAAGTGTCCAACTTCGAC 829
 QY 832 CCAAGGTAACCGGTTTGTATCCGATTAATCCGACACCGGTATCAACCGGTAT 891
 DB 830 CCATGCTGAACCTTTGCTGCGACATATGCTTAACACCGAGTAAACCGAA 889
 QY 892 TTGCTCTGATTTGTCGACGTTTGTCTCTCCCTCCAGTTCAATTAATTCGAC 951
 DB 890 ACCAATGATTTTGTGACGTTCTCTCTCCCGAACCCTGATCAATATACAGAC 949
 QY 952 CCGGACATATCCGCTTCTTGGGATATATTCAGAGAACAGACAGTTGGGTC 1011
 DB 950 TCCGACATGATCGTGTACTTGGAAATGATATTCAGAGAACAGACAGTTGGGTC 1009
 QY 1012 TTATCAGATGATCTCGTAGAGTGTCTTCATCAGATATGCAATCAACGATACA 1071
 DB 1010 TTGATAGTGAATCTCGGAGATGCGGCTTCATCTCATGTGACGTCGAAAGTTCAA 1069
 QY 1072 AAGCAGTGAATCAAGTATCGGAATCAAGACCTTAATGATGATCTTGGCTCA 1131
 DB 1070 GAGAGCTAGATGAGTTGTCGAAAGACGCTCCGACAGAGTACGTCGACGTG 1129
 QY 1132 CTTCATATCTAAAGCTGTGTGAAGATATTTAGGCTTCACTCCAGGCCACTT 1191
 DB 1130 ATGACGATCTAACAGAGGTGTGAAGAGGTGTGCGGTGACACCCCGGCCACTT 1189
 QY 1192 CTATCATGAGCGCCGTTGGCCATTAACAGACGATGATGATGATCTTGTTCGGCA 1251
 DB 1190 CTATCATGAGCGCCGTTGTCCATCAATGATGACGATTAATGATGATGATGATG 1249
 QY 1252 GGGACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1311
 DB 1250 GGGACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1309
 QY 1312 CTATTGAGTTAAACCTGAGAGTTGTGCAAAAAGAGTGAAGTTTGGCTT 1371
 DB 1310 CCATCTGAAATTAAGCCCGAGAGTTGTCACTCGGAGTGAAGTGCAGATTTTCATA 1369
 QY 1372 CTGAGGTCGATTTGAGACTGACCTTGGGTCGAGTGTGAGTTTGGCGGAG 1431
 DB 1370 CTGAGGTCGATTTGAGACTGATGATGATGATGATGATGATGATGATGATGATG 1429
 QY 1432 AATCTGTTTACTACCGTTATGTTTGAAGGCGATGATGATGATGATGATGATGAT 1491
 DB 1430 ACTCTGATGAGGCTACCGTTATGTTTGAAGGCGATGATGATGATGATGATGATGAT 1489
 QY 1492 GAGACGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1551
 DB 1490 GTACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1549
 QY 1552 GCTAATCT 1593

DB 1550 GCTAATCT 1591
 RESULT 5
 ABR23239
 ID ABR23239 standard; DNA; 1764 BP.
 XX
 AC ABR23239;
 XX
 DT 01-MAY-2003 (first entry)
 XX
 DE Seed development embryo/endosperm size alteration DNA SEQ ID 20.
 XX
 KW Cytochrome P450; embryo; endosperm; size; seed development;
 XX antisense inhibition; co-suppression; transformed plant; gene; ds.
 XX
 OS Glycine max.
 XX
 PN WO20029063-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 04-JUN-2002; 2002WO-US17562.
 XX
 PR 05-JUN-2001; 2001US-295921P.
 PR 28-NOV-2001; 2001US-334317P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Cahoon RE, Heppard EP, Nagasawa N, Sakai H;
 XX
 DR MPI: 2003-229292/22.
 XX
 F-PSDB; ABR26683.
 XX
 PT New isolated nucleic acid fragment for antisense inhibition or
 PT co-suppression of a cytochrome P450 polypeptide associated with
 PT controlling embryo/endosperm size during seed development in plants
 XX
 PS Claim 38; Page 86-87, 137pp; English.
 XX
 CC The invention relates to an isolated nucleic acid fragment comprising a
 CC promoter consisting essentially of a sequence having 2000, 2022, 8300 or
 CC 17201 base pairs fully defined in the specification, or their similar and
 CC functionally equivalent fragment or subfragment. The nucleic acid encodes
 CC a cytochrome P450 polypeptide associated with controlling embryo/
 CC endosperm size during seed development. The nucleic acid fragment is
 CC useful in antisense inhibition or co-suppression of a cytochrome P450
 CC polypeptide associated with controlling embryo/endosperm size during seed
 CC development in a transformed plant. This polynucleotide sequence
 CC represents a gene relating to the alteration of embryo/endosperm size
 CC during seed development of the invention.
 CC
 CC Sequence 1764 BP; 424 A; 459 C; 440 G; 441 T; 0 other;
 SQ
 Query Match 37.6%; Score 599.6; DB 25; Length 1764;
 Best Local Similarity 65.0%; Pred. No. 6.8e-193; Indels 42; Gaps 7;
 Matches 1014; Conservative 0; Mismatches 504;

QY 64 CAACCAACCTTCCCTTCT 120
 DB 133 CAAGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 192
 QY 121 CTCTTCTTATGACCTATTCGCGTGAAGCTGTGCGGGAATTAACCTTTGG-----C 174
 DB 193 TTTCTTACTGTCTCACCTGT 252
 QY 175 CGGTAATATCCGTTTCTATCAAAACCGGAAAG-----TTATTCGCGGTCA 222
 DB 253 TGGAAAAAACCACTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 312
 QY 223 AAGGCTTCCCTTGTGTAAGCATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 282


```

Db 313 AGAGGCTACCCCTTTCATTTGGAGATGAGTCTCATGATCC---CTGCGACACACCGT 369
Qy 283 ATGGCTGATGACGTGAGAAATTGCGAGCCAGAGGCTCATGGCTTTTCAGCTTAGAGAG 342
Db 370 ATTGCTGCGGCGGGGAGAGATGCAAGCCACAGGCTCATGGCTTTTCAGTGGTGA 429
Qy 343 ACTGCGGTGATCGTCACTGCAATCCGACGTAGCGAAGAGATTTGAAATAGCCGGT 402
Db 430 ACAAGCGCCATAGTAAAGTGCACACCCGATGCTGCTAAAGATTTCTCATATGTTCCACT 489
Qy 403 TTGGCTGATGACCGGTTTAAAGATGGCTTCACTCATGATGTTTAAAGAGCAATGGT 462
Db 490 TTGCTGATGCTCCCATTAAGAGATCAGCTTACAGCTCATGTTCAACCGCGCATCGG 549
Qy 463 TTGCGACACACGCGTGTCTTACGCGCAACGCTTGCCTGATCGCTTGAACCATCTCTT 522
Db 550 TTGCGCCCTTAACGCGCTCTACTGCGCTACCTCGCGCATCGCGCATCGCACCTTTC 609
Qy 523 AGTACAAACAAATCAGAGAGCGGACGCAAGCAAGTGAATCTCAAGCCAGATGTT 582
Db 610 TGCCCAACAAATCAAGCTCTCGAGCTCAGCGCTAAATCGCCCAATGACA 669
Qy 583 GAGTTCTTGAAGAAAGAGATGATGACAAACCTGTTTGTGAGTGTCTTAAACG 642
Db 670 AACTCATTCGAAATCACCGTTGACGCGGCTTTCGAAATCGCAGCGCTCAAGAGA 729
Qy 643 GCGTGTCTTAAACAATGATGCTCTGATTCGCAACAAGATGAGCTTGAAGAAAC 702
Db 730 GCGTCACTGACCAACATGATGATGCTGCTTGAACAAGTCAACCTTGAAGAGATA 789
Qy 703 CA-----TGTTGAGTACGTGAATGCTGCAAGAGTGAATGATTTGCTCGAAG 753
Db 790 AACACCGAATGAGAGAGCTATCCATGTTGATGAACAAGGCTATGATGTTGGGAC 849
Qy 754 TTGAATGAGTGAATCATTCTCTGCTTGGCTATCGAGATTTGATCTCAAAAGCTCGGTCT 813
Db 850 CTTAATGCGGAGACCAATCCCTTCTTGAAACCTTGAACCTTGAACGAAATCGGTTTC 909
Qy 814 AGATGTTCCACATCGTACCAAGGTAACCGGTTTGTGATCCCGATTTATATCCGAAC 873
Db 910 ACGTCTCTCAATTAATGCTCTCAAGTGAACCGGTTCTGTTGATCAATCGCGACAC 969
Qy 874 CGTATCAACACCGGATTTGCTGCTGATTTCTGCAAGTTCCTGCTCTCCATGCT 933
Db 970 CAGGCGCACACACCAACCAACACCGGATTTCTGATGTTTGTCTCTCTCCCAAGGT 1029
Qy 934 TCAGTAAATTAATCCGACCGGACATATGCGCGTCTTTGGAGATGATATCAAGAGA 993
Db 1030 CCGGTAATTAATGCTCACTCCGACATGATGCTGCTCTCTGGAATTAATTTAAGGGG 1089
Qy 994 ACAAGACAGATGCGGTCTTAATGAGTGAATCTCTGCTAGAGTGTCTTCAATCCAGAT 1053
Db 1090 ACCGACAGGTGCGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1149
Qy 1054 ATGCATCAACGCTCAACAAAGAGTGAATCAAGATGCGGAAATCAAGAGCCCTGAT 1113
Db 1150 GTGCAAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206
Qy 1114 GATCTGAATGCTGCTTCACTTCCATATGATGATGATGATGATGATGATGATGATGATGAT 1173
Db 1207 GAGGAGGTGCTGCGCGGACGCGCTATCTGCGCGGCTGATGATGATGATGATGATGATGAT 1266
Qy 1174 CATCTCCAGGCGCTTCTATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1233
Db 1267 CACCGCGCGCGCTGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
Qy 1234 GGTGCTTGTTCGCGCAGGAGCAACAGCAATGATGAACATGATGATGATGATGATGATGAT 1293
Db 1327 GGGTATCAAGTGTGCGGAGCAACCGCTATGATGATGATGATGATGATGATGATGATGAT 1386
Qy 1294 CCAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1353
Db 1387 CCGGAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1443

```

```

Qy 1354 GAGTGAAGTTTTCGTTCTTGGGTCGGATTTGAGACTTGTGACCTTTCGGGTCGGTCT 1413
Db 1444 ---AACGAGTTTCTGTTTCGGGTCGGATCTGAGACTGCTCCATTCGTTCCGGTCCG 1500
Qy 1414 CGGATTTGCGCGGAGAAATCTTGTGTTTAACTACCGTTATGTTTGAAGCGGATGATG 1473
Db 1501 AGAACATGCGCGGAGAACTTGGGTTGAGACCGGTAACCTTCTGGGTGGCTT 1560
Qy 1474 TTACATGATTTGAATGGGACCGTCCGATGATGACCGGTTGATCTGAGAAACTG 1533
Db 1561 TTGATGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Qy 1534 AGGCTTCTGCGAGATGCTATCTCTTCTGCTTAAATTCGCGGATGAGCGCATGTA 1593
Db 1621 AGCTCTGCTGTAATGCTTAACCACTCATTTTAAAGTTGCGCCCTAGGATGATTA 1680

RESULT 6
ABT23241
ID ABT23241 standard; DNA; 1905 BP.
XX
AC ABT23241;
XX
DT 01-MAY-2003 (first entry)
XX
DE Seed development embryo/endosperm size alteration DNA SEQ ID 24.
XX
KW Cytochrome P450; embryo; endosperm; size; seed development;
XX antisense inhibition; co-suppression; transformed plant; gene; ds.
XX
OS Glycine max.
XX
PN MO200299063-A2.
XX
PD 12-DEC-2002.
XX
PF 04-JUN-2002; 2002MO-US17562.
XX
PR 05-JUN-2001; 2001US-295921P.
XX PR 28-NOV-2001; 2001US-334317P.
XX
PA (DUPC) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon RE, Heppard EP, Nagasawa N, Sakai H;
XX
DR WPI; 2003-229292/22.
XX
PT P-PSDB; ABT26685.
XX
PT New isolated nucleic acid fragment for antisense inhibition or
XX co-suppression of a cytochrome P450 polypeptide associated with
XX controlling embryo/endosperm size during seed development in plants
XX
PS Claim 38; Page 90; 137pp; English.
XX
CC The invention relates to an isolated nucleic acid fragment comprising a
XX promoter consisting essentially of a sequence having 2000, 2022, 8300 or
XX 1201 base pairs fully defined in the specification, or their similar and
XX functionally equivalent fragment or subfragment. The nucleic acid encodes
XX a cytochrome P450 polypeptide development. The nucleic acid fragment is
XX endosperm size during seed development. The nucleic acid fragment is
XX useful in antisense inhibition or co-suppression of a cytochrome P450
XX polypeptide associated with controlling embryo/endosperm size during seed
XX development in a transformed plant. This polynucleotide sequence
XX represents a gene relating to the alteration of embryo/endosperm size
XX during seed development of the invention.
XX
SQ Sequence 1905 BP; 488 A; 498 C; 463 G; 456 T; 0 other;
XX

Query Match 36.8%; Score 585.8; DB 25; Length 1905;
Best Local Similarity 66.1%; Pred. No. 3.6e-188;
Matches 934; Conservative 0; Mismatches 452; Indels 27; Gaps 5;

```

QY 193 TACAAACCGGAAAGCTATATCCCGGTCCAAAAGCTTCCCTTGGTGGAGCATGTCA 252
 Db 294 TCCAAACACCTTAAATGATTCGGGTCCAAAGGCTATCTTTCAATGGAAAGCATGAGC 353
 QY 253 CTCATGTCAAGCACTAGCTCACCAGCAATGCGTATGACAGTGAATTCGAGCC 312
 Db 354 CTCATGACATCC---CTTGACACCAACCGATTTGCTGCGCTGTCTCAACATGCAAGCC 410
 QY 313 AAGAGGTCAATGCTTCACTTGAAGAGACTGCGGTGATGCTCACTGCAATCCGAC 372
 Db 411 ACCAGGCTCATGCGCTTCCATGAGGTGACACCGGTGATGCTCACTGCAATCCGAC 470
 QY 373 GTAGCGAAAGATTCGAAATAGCCGGTCTTTGTGATGACAGCGGTAAAGATGAGCT 432
 Db 471 GTGCGCAAGAGATTCCTTAACAGCTCGTCTTCCGATGCTCCATTAAGAGATGAGC 530
 QY 433 TACTCACTGATGTTTAAAGAGCAATGCTTTGACACACGCTGTTTACTGCGAAG 492
 Db 531 TACAGCTCATGTTCAACCGCGCATCGCTTGGCCCTTACCGCTTTACTGCGCAC 590
 QY 493 CTTCGCGGTATGCTTGAACCATCTCTTAATGATCAAAACCAATCAAGAGCGAGAG 552
 Db 591 CTCGCGCGCATGCGCGCACGACCTCTTCTGCGCCCAACCAATCAAGGCTGAGCTC 650
 QY 553 CAACGACGATGATCTCAGACCAAGATGATGATTTCTTGAACCAAGATGATGAA 612
 Db 651 CAGCGCGCCCAATTCGCGCGCACGATGACCACTCGTCCGAAAC---GCCGCGC 704
 QY 613 CCTGTTTCTTGTGATGCTTAAACCGCGCTGCTTAAACCAATGATGCTGCTGTA 672
 Db 705 GGTTCGGAATCCGACGCTTCTCAAGAGGCTCGCTCAACCAATGATGCTGCTG 764
 QY 673 TTGCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 723
 Db 765 TTGGAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 824
 QY 724 GTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 783
 Db 825 GTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 884
 QY 784 TCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 843
 Db 885 AAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 944
 QY 844 CGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 903
 Db 945 CGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1004
 QY 904 TTGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 963
 Db 1005 TTGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1064
 QY 964 GCCGTTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1023
 Db 1065 GCTGCTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1124
 QY 1024 ATCTCGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1083
 Db 1125 ATCTCGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1184
 QY 1084 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 Db 1185 GCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1244
 QY 1141 CTAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
 Db 1245 CTTCTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1304
 QY 1201 GCCGTTTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
 Db 1305 GCCGCTTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1364
 QY 1261 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320

Db 1365 GCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1424
 QY 1321 TTAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
 Db 1425 TTCAAGCCCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1478
 QY 1381 GATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
 Db 1479 GATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1538
 QY 1441 TTATCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
 Db 1539 TTGAGCACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1598
 QY 1501 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
 Db 1599 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1658
 QY 1561 CTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1593
 Db 1659 CTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1691
 RESULT 7
 AAX55600
 ID AAX55600 standard; DNA; 1799 BP.
 XX
 AC AAX55600;
 DT 26-JUL-1999 (first entry)
 XX
 DE Clone 040 ovule-specific gene encoding a cytochrome P450 monooxygenase.
 XX
 KW Ovule-specific gene; genetic engineering; ovule; nutritional value;
 XX fruit; seed; crop plant; cytochrome P450 monooxygenase; ss.
 OS Phalaenopsis sp.
 PN US5907082-A.
 XX
 PD 25-MAY-1999.
 XX
 PF 17-NOV-1995; 95US-0560398.
 XX
 PR 17-NOV-1995; 95US-0560398.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Nadeau J, O'Neill S;
 XX MPI, 1999-337265/28.
 DR P-PsDB; AAT13438.
 XX
 PT Ovule-specific gene expression - useful for targeting gene
 XX expression to ovules or for modulating ovule development
 XX
 PS Examples; Columns 33-36; 32pp; English.
 CC The invention provides ovule-specific genes which are useful in
 CC genetically engineering plants. The use of specific promoters from the
 CC genes of the invention allows directed expression (or inhibition) of
 CC desired heterologous genes in ovules e.g. to increase the nutritional
 CC value of fruit or seed, or to inhibit the development of ovules. The
 CC polynucleotide sequences (AAX55599, AAX55601-604) are expressed uniquely
 CC during ovule development and as such they play a role in regulating
 CC development and establishing the specialised sequences are able to be
 CC ovule. The isolated genes or their promoter sequences are used to be
 CC used to produce improved crop plants. The present sequence represents a
 CC clone 040 ovule-specific gene encoding a cytochrome P450 monooxygenase
 CC from the pollen tubes of Phalaenopsis.
 CC
 CC Sequence 1799 BP; 456 A; 362 C; 428 G; 548 T; 5 other:

Query Match 22.3%; Score 355; DB 20; Length 1799;
 Best Local Similarity 58.0%; Pred. No. 1,6e-109;
 Matches 744; Conservative 0; Mismatches 500; Indels 39; Gaps 5;

QY 304 TTGGAGCCCAAGAGCTCATGCTTTGAGCTTTAGAGAGACTGCGTGAATCGACAGTCG 363
 DB 273 TTGACGCGCCCTCGGCTCATGCTTTCTCGGTGGGCTCACTGCTTCAATCGTTTAAAGC 332
 QY 364 AATCCGACGTAGGAGAAAGAGATCTGAATAGCCCGGTTTGTGATGACCGGTAA 423
 DB 333 CACCCGAAACCGCAAAAGAGATCTTCAAGCCGCTTGGCTGATGGGCCATTA 392
 QY 424 GAATCGGCTTACTCATGATGTTTACAGAGCAATGGTTTGCACCAACGCTTTAC 483
 DB 393 GAATCAGCATACGAACTTCTGTTTATCGGCTATGGGTTTTCGCCATTTGGGGATTAC 452
 QY 484 TTGGGAAAGCTTGGCCGATGCGTTGCAACCATCTTTAGTCAAAAACAATCAGAGA 543
 DB 453 TGGGAAACCTGAGAGAGATTTCTCCACATATCTTTTCACTCGCGCGAGTTGATCG 512
 QY 544 GCCGAGACGCAACGACGATGATCTCAAGCCAGTGTGAGTTTCTTGAAGAAACGA-- 601
 DB 513 TTGAGAGACGACGAGTGAATGGCGAAGATGTTGGGGATATGAAAGAAATGATG 572
 QY 602 -GTAGTACGAAACCTGTTTGTTCGTGAGTGTCTTAAACGGGCTGCTTAAACAATG 660
 DB 573 GAGGAAATGAGTGTGAGAGTGAAGAGAAAGTTGCACTACGGGCTTTGAATACATC 632
 QY 661 ATGTCTCTGTATTGCGACAAAGATGAGCTTTGAAAAAACATGTTGAGTTACGTGA 720
 DB 633 ATGTGACTGTTTTGGGAAAAAGTTGATTTGCAAAAGATGAGGGGTTGAGCTTGAG 692
 QY 721 ATGGTC--GAAAGAGTTATGATTTGTCGGAACGTTGATGATGATGATGATGCTCT 777
 DB 693 TTGATCCTTAAAGAGAGATGAGTACTTGGGATCTTCAACTGAGGATCATTTGCT 752
 QY 778 TGGCTATCGAGTTGATCTCTCAAGACTCCGGTCTAGATGTTCCACTGTAACAAG 837
 DB 753 CTTTGGAGATGTTGATTTGGAAGTGTGAGAGAAATGATCAGAAACATTTGGCTAAG 812
 QY 838 GTAAACCGGTTGTATCCCGGATTTATCCGAAACACCGTAATCAACC----- 885
 DB 813 GTCAATGTATTTGTAGAGAGATCATAGACGACGATTAAGAGAGCCGAGGTAGGG 872
 QY 886 -----GGTGAATTCCTCGTGAATTTGTCGACGTTTGTCTCCCTCATGCTTCAAT 939
 DB 873 ATTGATGAGGAGTGAAGTGAAGATTTTGTGATGTCCTTCTTGTTGAGAGAGAAAGAT 932
 QY 940 AATATTATCGAACCCGACATTAATGCGCGCTTTTGGGAGATGATTTCAAGAGAAACAGAC 999
 DB 933 AAGACTCTCAGATCTGATATGATGTCGACAGTTCTTTGGGAAATGATCTTTAGAGAACTGAT 992
 QY 1000 AAGATTGGCGCTTATATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1059
 DB 993 ACTGTTGCCATCTTATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1052
 QY 1060 TCAACGGTACAAAACGAGCTGATCAAGTATGCGGAAATCAAGAGCCCTAGATGATCT 1119
 DB 1053 TCGAAGGCAACAAGTTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1112
 QY 1120 GATTTGGCTTCACTTCAATATCAAGCGCTGTGAGAAAGATATGAGGCTTATCT 1179
 DB 1113 GATATCAACGATCTTATATCTTCAATATATGTAAGAAACCTTCAATGATCT 1172
 QY 1180 CCAAGGCCATTCTATCATGAGGCGCTTTGGCCATTAACAGACAGATGATGATGAT 1239
 DB 1173 CCGGAGCTCTATTTGATGAGGCTCGCTAGTATCAATGAGATGATGATGATGATGAT 1232
 QY 1240 CTTGTTCCGAGAGGACCAAGATGAGTGAACATGAGGCGGATTCGATATCAACAC 1299
 DB 1233 ATGATTCCTGCTGGAGACATGAGTGAACATGAGTGAACATGAGTGAACATGAGTGAAC 1292

QY 1300 GTGTGGTGTATCTTTGAGATTAAACCTGAGAGGTTCTGGCAAAAGAGTGAAGTG 1359
 DB 1293 AACTGGGCTGAGCTTAACAATTCATCTGATGATTCATGATGAAGT----- 1344
 QY 1360 GAGTTTTCGTTCTTGGGTCGAGATTGAGACTTTCGACCTTTCGGGTCGGTCTGAT 1419
 DB 1345 ----TCAATATCTTGGTTCGATTTAAGTTGGCACCTTGGCTCCGGTAAAGATT 1400
 QY 1420 TGGCCCGGAAAGAAATCTTGTATTACATACGTTATGTTTGAACGGGATGATGATCAT 1479
 DB 1401 TGGCCCTGGCAAAACGATGGCAATGAGCTGACATCTTGTGTTGGCTGATGCTGAAA 1460
 QY 1480 GAGTTTGAATGGGAGACGCTCCGATGTAACGCGCTTACTATCTGAGAACTGAGGCT 1539
 DB 1461 AGCTTCAATTCCTTCTCTC--GAGAAATGCTGATGATTTGCTGAGTGCCTAAGATG 1517
 QY 1540 TCTTCCGATGATGCTAATCTCT 1562
 DB 1518 TCTCTGAGATGAGAAATCTTT 1540

RESULT 8
 ABT23246
 ID ABT23246 standard; DNA; 1930 BP.
 XX
 AC ABT23246;
 XX
 DT 01-MAY-2003 (first entry)
 XX
 DE Seed development embryo/endosperm size alteration DNA SEQ ID 34.
 XX
 XX Cytochrome P450; embryo; endosperm; size; seed development;
 XX antisense inhibition; co-suppression; transformed plant; gene; ds.
 XX
 OS Aquilegia vulgaris.
 XX
 PN WO200299063-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 04-JUN-2002; 2002WO-US17562.
 XX
 PR 05-JUN-2001; 2001US-295921P.
 XX
 PR 28-NOV-2001; 2001US-334317P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Cahoon RE, Heppard BP, Nagasawa N, Sakai H;
 XX
 DR WPI; 2003-229292/22.
 XX
 DR P-PSDB; ABJ26650.
 XX
 PT New isolated nucleic acid fragment for antisense inhibition or
 XX co-suppression of a cytochrome P450 polypeptide associated with
 XX controlling embryo/endosperm size during seed development in plants
 XX
 PS Claim 38; Page 97-98; 137p; English.
 XX
 XX The invention relates to an isolated nucleic acid fragment comprising a
 XX promoter consisting essentially of a sequence having 2000, 2022, 8300 or
 XX 17201 base pairs fully defined in the specification. The nucleic acid encodes
 XX functionally equivalent fragment or subfragment. The nucleic acid encodes
 XX a cytochrome P450 polypeptide associated with controlling embryo/
 XX endosperm size during seed development. The nucleic acid fragment is
 XX useful in antisense inhibition or co-suppression of a cytochrome P450
 XX polypeptide associated with controlling embryo/endosperm size during seed
 XX development in a transformed plant. This polynucleotide sequence
 XX represents a gene relating to the alteration of embryo/endosperm size
 XX during seed development of the invention.

Query Match 22.1%; Score 352; DB 25; Length 1930;

Best Local Similarity 56.3%; Pred. No. 1.8e-108;
Matches 786; Conservative 0; Mismatches 565; Indels 45; Gaps 5

197	AAACCGGAAACGTTATCCCGATCCAAAGAGCTCCCTTGTTGAGAGATGCACTCA	256
213	AMGCTAAACCGCATTCCTGACCACTCGGGTTCTTTATTTGGGTTGGTCTTGGCTT	272
257	TGTCAGACACTCTAGCTCCACCGAATCGCTATGACGTCGAAATTTGGAGCCAGA	316
273	TTACTGGTTCTACTACTCATAGAGTTTGTAGCAATCTTGCTAAACCTTTAAAGCTATTC	332
317	GGCTCATGGCTTTGAGCTTAGAGAGAGACTCGCGTATCGTCACGTCGATCCGACGTAG	376
333	CTTTAAATGGCTTTTCTGGTTGGTTTACTCGTTTATCATATCAAGTTGTCGTATACAG	392
377	CGAAAGAGATTCTGAATATAGCCCGGTTTTTGTGTATGACCGGTTAAAGATTCGGCTACT	436
393	CAAAAGAGATTCTTAATATGATTCCTTTGTGTATGACCGGTTAAAGATTCGTATAG	452
437	CACGTAGATTAAACAGACCAATTTGGTTCACACACAGGTTTACTGGCGAACGCTTC	486
453	AACTTTGTTCACAGAGCAATGGGTTTGTGCTCTTTTGGATATTTGAGAGAACTGTA	512
497	GGCGTATCGCTCGAACCACTCTCTTTAGTCAAAACAAATCAGAAAGCCGAGACGCAAC	556
513	GAAGAAATCTCAGTACCATTATTCAGTCCAAAGAGATTAACCGGTTTGTCTACATTC	572
557	GACAGATGATCTCAGCCAGATGGTTGATGTTCTTGAAAAACAGATGAG--TTACGAC	613
573	GAATGAAATATGAGAAAAAATGATATATGATGATTAATGTCAAATGGGGTTAATAGGG	632
614	CCGTGTTTGTTCGAGAGTTGCTTAAACCGCGTGCCTTAACAATGATGTCCTGAT	673
633	AAGTTGAAGTTAAAGGGTATTACACTTTGGGCTCTTAAACAATGATATGATGATGTTT	692
674	TCGGAACAAGATATGAGCTTGAAAAAAACATGTTGAGTTACGT-----GAATGG	724
693	TTGGAGAGCTTTATGATTTTAAACAACTTAATGATGATGGGTTTAACTTGAAAGATGG	752
725	TCGAAGAAGTTATGATTTGCTCGAAGCGTGAATTTGACATGATCACCTTCCTGGCTAT	784
753	TGAGTGAAGGATATGAGTTGCTGGGATTTTAACTGAGAGATCACTTTCCTTTATGG	812
785	CGAGATTGATCTCTCAAAAGACTCCGGCTAGATGTTCCACACTGTACCAAAAGTTAAAC	844
813	GCTGGTGGATTTGGACAGAGATTAAGGAAGAACAGAGTGTGTTCTTAAGGTGATAA	872
845	GGTTGTATCCCGGATTTATCTCGAACAACCGTTAATCAACCGGT-----	888
873	TTTTTGTGGAAAAATTTATGGAAGAACACAAACAGAAAGATTAATGATGTTTGGGTC	932
889	--GATTGCGCTCGATTTGCTGCACGTTTGTCTCCCTCCATGCTTCAGATTAATAT	946
933	AAGATATGTTGGTACCTTTGTATGTCCTTGTGATTTGGAAGAAAGAAATATGATCTCA	992
947	CCGACCCCGGACATAATCCCGGCTCTTTGGAGATGATATTTGAGAGAACACACACAGTTG	1006
993	GTGACTGTGACATGATCTGTTCTTTGGAAATGATCTTTAAGGGGACACACACAGTAG	1052
1007	CGGTTTAAATGCAATGATCTCTCGTAGAGATGCTTCATCCAGATTTGCAATCAACGG	1066
1053	CAATCTCTTAAGAGGATCTTTGGAAGAAATGGCCCTACATCCAGATTAATCAAGCAAAAG	1112
1067	TACAAACAGAGCTGATCAAGATAGTCGGGAATCAAGACCCCTAGATGAATCTGACTTGG	1126
1113	CCCAATCTGAATATGACACTGTGCTTGGCACTTAATCGACTATGATCTGACTTAC	1172
1127	CTTCACTTCCAAATCTTAACGGGTGGTGAAGAAGATATGAGGCTTATCTCTCAAGGCC	1186
1173	CCAACTTCTTATCTCCAGAGATGATGAAGAAATCTTTAAGGATGACCTTCTTGACC	1232
1187	CACCTTCTCATGGGCGGTTTGGCAATACAGACAGATCGTTGATGATGATGCTTGTGTT	1246

Db	1233	CCCTCTTGTGTGGGACGACCTAGCTATCCATGATGTCCCATATGGAAGAACTTTATCC	1292
QY	1247	CGGCAAGGACACACAGCATAGTGAACATGTGTGGCCGCTATGGCANTGATCCACAGTGTGGG	1306
Db	1293	CAGCTGGACACTACGCTATGGTGAATGTGGGCAATCACCTCAGTAAGAAAGTATTTGGT	1352
QY	1307	TTGATCCCTTTGGAGTTTAAACCTCAGAGGTTCTGTGGCAAAAGAGGTGAGGTGAGTTT	1366
Db	1353	CTGAGCCAAATGAATTTAAACCAGAGGATTCATGCATGAAAGT-----TGA	1400
QY	1367	CGGTTCTTGTGGTCGGAATTTGAGACTTGCACCTTTGGGTCGGGTCGTGGATTTGGCCCG	1426
Db	1401	GCATTATGTGGGTCGATCTGAGGTTGGCACTTTGGGTCGGAAGAGGGTTTGTCTGT	1460
QY	1427	GGAGAAATCTTGGTTTACTACCGCTTATGTTTGGAGAGGGCGATGATGTTACATAGTTTG	1486
Db	1461	GAAAGGCTTGGGCTATAGGCTACTGTGACGCTATGGTGGGTCACTTACTTCAAAAGTTTCA	1520
QY	1487	AATGGGGAACGCTCCGATGTGTAAGGGCGGTGACTATCTGAGAACTGAGGCTTCTTTCG	1546
Db	1521	AATGGGTTCTTCTGAAAGT---GCTGGAACCTTGCTGAGTGTCTTAATCTGTCTCTGG	1577
QY	1547	AGATGGCTAATCTCT	1562
Db	1578	AAATGAGAAAGCTTT	1593

CC	polypeptide associated with controlling embryo/endosperm size during seed development in a transformed plant; this polynucleotide sequence
CC	useful in antisense inhibition or co-suppression of a cytochrome P450
CC	endosperm size during seed development. The nucleic acid fragment is
CC	a cytochrome P450 polypeptide associated with controlling embryo/co-
CC	functional equivalent fragment or subfragment. The nucleic acid encodes
CC	17201 base pairs fully defined in the specification, or their similar
CC	promoter consisting essentially of a sequence having 2000, 2022, 8300 or
CC	The invention relates to an isolated nucleic acid fragment comprising a
PS	Claim 38; Page 88-89; 137pp; English.
XX	
XX	
PT	controlling embryo/endosperm size during seed development in plants -
PT	co-suppression of a cytochrome P450 polypeptide associated with
PR	New isolated nucleic acid fragment for antisense inhibition or
XX	(DUPO) DU PONT DE NEMOURS & CO E. I.
PA	
XX	
XX	
PI	Cahoon RE, Heppard EP, Nagasawa N, Sakai H;
XX	
XX	WPI; 2003-229292/22.
DR	P-PSDB; AB326684.
XX	
PF	04-JUN-2002; 2002WO-US17562.
XX	
XX	05-JUN-2001; 2001US-295921P.
PR	28-NOV-2001; 2001US-334317P.
XX	
XX	
PD	12-DEC-2002.
XX	
PN	WO200239063-A2.
XX	
OS	Glycine max.
XX	
KM	Cytochrome P450; embryo; endosperm; size; seed development; antisense inhibition; co-suppression; transformed plant; gene; ds.
DE	Seed development embryo/endosperm size alteration DNA SEQ ID 22.
DT	01-MAY-2003 (first entry)
XX	
AC	ABT23240;
XX	
XX	ABT23240 standard; DNA; 1934 BP.
ID	ABT23240
RESULT 9	

CC represents a gene relating to the alteration of embryo/endosperm size
 CC during seed development of the invention.

XX Sequence 1934 BP; 412 A; 386 C; 547 G; 589 T; 0 other;

Query Match 19.9%; Score 316.6; DB 25; Length 1934;
 Best Local Similarity 55.1%; Pred. No. 2.1e-96;
 Matches 750; Conservative 0; Mismatches 569; Indels 42; Gaps 5;

254 TCATGTCAAGCACTCTAGCTCAACGAGATGCTGATGACGTGAAGAAATTGGAGCCA 313
 DB TCTTACTAGGCTCCACGCTCAGCGTCTTATCCAACTCGCCGTAATTAACCAACCGG 314
 QY AGAGGCTAGGCTTTAGCTAGAGAGACTCGCGATGCTGACGAGCAATCCGAGC 373
 DB 314 AGAGGCTAGGCTTTAGCTAGAGAGACTCGCGATGCTGACGAGCAATCCGAGC 373
 DB 315 AAAAAGCTATGAGCTTTCTCCATCGCTTAAACCGCTTCTCATCTCCAGGAAACCGAGA 374
 QY 374 TAGCGAAAGATTCGAAATACCGCGTTTCTGATGACCGGTTAAAGATCGGCTT 433
 DB 375 CCGCTAAGAGATTCGCGAGAGCCCGAGTTTCTGATGAGCCGCTGAAGAAATCCGCT 434
 QY 434 ACTCATGATGTTTACAGAGCAATGCTTTTGAACACAGCGTTTACTGCGAGAGC 493
 DB 435 ATGAGCTTCTCTCCACCGCGAATGGGTTTTCACCGTATGGGAGTACTGAGGAATT 494
 QY 494 TTCGCCGTATGCTTCGAAACCATCTCTTATGACAAACAAATCAGAAAGCCGAGACG 553
 DB 495 TGAGGAGATTCGAGCCCTACATCTCTCCCGAAGAAATCAGCGCTCTGAATCCT 554
 QY 554 AACGACGATATCTCAAGCCAGATGCTGATGCTTCTGAAAACAGATGCT---AACG 610
 DB 555 TCAGAGGAGGATGATTAATAAAGTTGAACAGTTAAACAAACCTGATGAGAAC 614
 QY 611 AACCTGTTTGTTCGTAGTGTCTTAAACGCGCTCGCTTACAGCATGATGCTCTG 670
 DB 615 AACATGTTAGGTTAAGAAATTTACACTTATGCTGTAACATGATGATGAGCG 674
 QY 671 TATTTCGACAGATGATGAGCTT---GAAAAAACATGTTGATGATGATGATGCTG 727
 DB 675 TGTTCGATGATCTTATGATGATTTAAGAGGTTGAGGTTGAGCTTGAAGGTTTGGTGA 734
 QY 728 AAGAGGTTATGATTTGCTCGGAGCGTTGAATGAGTATGATCAGCTTCTGCTATCG 787
 DB 735 GTGAAGGATGATGATGATGATGATTTTAACTGAGAGCAATTTCCGCTTTTGGGGT 794
 QY 788 AGTTGATCTCAAGACTCCGCTGATGATGTTTCAACTGTCACAAAGGTTAAACGGT 847
 DB 795 GATTGATTTGAGGAGGTGAGAGAGAGGTGATGATGATGATGATGATGATGATGAT 854
 QY 848 TTGATCCCGGATTTATCCGACACCGTAAATCAACCGG----- 887
 DB 855 TTGTTGAGAGGTTTAAAGAGCAATGAGGTGAAGAGGAGGAGGAGGAGGAGGAGG 914
 QY 888 -TGATTTCCTGATTTGCTGACGCTTTGCTCTCCCTTCATGCTTCAATTAATAT 946
 DB 915 ATGAAGGAACTGGGATTTTGTGATGTTTGTGATTTGAGAGAGAGAGAGAGGCTCA 974
 QY 947 CCGACCCGAGATTAATCCGCTTTTGGAGATGATTAATGAGAGAGAGAGAGAGAGT 1006
 DB 975 GTGAAGCTGATGATGATGATGATGATTTTGGGAAATATTTTGGGAACTGACAGCGT 1034
 QY 1007 CGGCTTAAATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066
 DB 1035 CAATTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1094
 QY 1067 TCAAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1126
 DB 1095 CACAGGCGAAATTAATCTTCTTGGGATCTTCCAGGCGCTTACCGAAGAGAGATTC 1134
 QY 1127 CTTCACTTCATATCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186
 DB 1155 CCAAGCTGCTACCTTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1214

QY 1187 CACTTCTATCATGAGGCGGCTTTGGCCATACAGACAGCATGATGATG---TCGCTTGG 1243
 DB 1215 CCTTACTCTCTGAGGCTGAGCTTCTGTCAGACAGATTAACCTGAGCGGCAAGCATGGA 1274
 QY 1244 TTCCGAGAGGACACAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1303
 DB 1275 TTCCCAAGGCGACACCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1334
 QY 1304 GGGTGTATCTTTGAGTTTAACTGAGAGGTTGATGATGATGATGATGATGATGATGATGAT 1363
 DB 1335 GGGCGAGCGCGAAGATTTAGGCGCGAGCGGTTGT-----GAGAGAGATG 1382
 QY 1364 TTTCGATCTTGGGTGAGATTTGAGATGACCTTTGCGGTGCGGTGCTGATGATGATGATGAT 1423
 DB 1383 TGAGCATATAGGAGCTGATTTGAGGTTGAGCACTTTGCGGTGCGGTGAGAGAGATGATGAT 1442
 QY 1424 CCGGAGAGATTTGATTTTACTACGCTTATGATGATGATGATGATGATGATGATGATGATGAT 1483
 DB 1443 CTGGAGAGGCTTGGTTGGCTGCTGATCATCTTTGCTGCTGATGATGATGATGATGATGAT 1502
 QY 1484 TTGAATGAGGAGCGCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1543
 DB 1503 TTCAATGAGGTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1562
 QY 1544 GCGAGATGCTATATCTCTCTCTCTTAATTTGCGCGCTTGG 1584
 DB 1563 TGAGATGAGAGAGCACTGTTGCAAGGCTGCTGAGG 1603

RESULT 10

ABT23249
 ID ABT23249 standard; DNA; 1687 BP.

ABT23249;

01-MAY-2003 (first entry)

DE Seed development embryo/endosperm size alteration DNA SEQ ID 40.

XX Cytochrome P450; embryo; endosperm; size; seed development;

XX antisense inhibition; co-suppression; transformed plant; gene; ds.

XX *Astroemeria caryophylla*.

XX WO200299063-A2.

XX 12-DEC-2002.

XX 04-JUN-2002; 2002WO-US17562.

XX 05-JUN-2001; 2001US-295921P.

XX 28-NOV-2001; 2001US-334317P.

XX (DUPO) DU PONT DE NEMOURS & CO E. I.

XX Cahoon RE, Heppard BP, Nagasawa N, Sakai H;

XX WPI; 2003-229292/22.

XX P-PSDB; ABJ2693.

XX New isolated nucleic acid fragment for antisense inhibition or

XX co-suppression of a cytochrome P450 polypeptide associated with

XX controlling embryo/endosperm size during seed development in plants

XX Claim 38; Page 101-102; 137pp; English.

XX The invention relates to an isolated nucleic acid fragment comprising a

XX promoter consisting essentially of a sequence having 2000, 2022, 8300 or

XX 11201 base pairs fully defined in the specification, or their similar and

XX functionally equivalent fragment or subfragment. The nucleic acid encodes

XX a cytochrome P450 polypeptide associated with controlling embryo/

XX endosperm size during seed development. The nucleic acid fragment is

XX useful in antisense inhibition or co-suppression of a cytochrome P450

CC polypeptide associated with controlling embryo/endosperm size during seed development in a transformed plant. This polynucleotide sequence represents a gene relating to the alteration of embryo/endosperm size during seed development of the invention.

XX Sequence 1687 BP; 246 A; 636 C; 456 G; 349 T; 0 other;

Query Match

Best Local Similarity 18.2%; Score 290.6; DB 25; Length 1687;

Matches 712; Conservative 0; Mismatches 579; Indels 24; Gaps 5;

257 TGTCAAGCACTAGCTACCGAGCAATGCTGATGAGTGAAGAAATTCGAGCCAA 316
 193 TCTCCGGCCCTCCGCGACCGCTCCCTGCGCGCGCTCTCTCTCCCTCCGCGCTCC 252
 317 GGCTCATGCTTTCAGCTTAGAGAGACTCGCTGATCTCAGTCAATCCCGAGTAG 376
 253 CCGTCT 312
 377 CGAAGAGATTCGATAGCCCGGTTTCTGATGACCGGTTAAAGATCGGCTTACT 436
 313 CCAAGGACATCTCTCCAGCTCCGCTTCGCGACCGCCCATCAAGATCCGCTTACG 372
 437 CACTGATGTTAAAGAGCAATGTTTTCACACACGCGTGTCTTTCGCGAAGCTTC 496
 373 GCTCTCTCTTCCACCGCGCAATGAGCTTCCGCTCTCTCTCTCTCTCTCTCTCTCT 432
 497 GCGGATCGCTTTCGAAACCATCTCTTTCGAAACAAATCAGAGAGCGGAGCAAC 556
 433 GCGGATCTCCGCGACCACTCTTTCAGCCCAAGCGCTCTCCGCTCCGCGCTCCG 492
 557 GAGGATGATCTCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 616
 493 GCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 552
 617 GTTCTTCTCTGAG--TTGCTTAAACGCGCTGCTTAAACATGATGATGATGATGAT 673
 553 AGGTCGATCAAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 612
 674 TCGGACAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730
 613 TCGGCGCGCTCTCAAGCTTCCGCAACCGGAGCGCTCTGAGCTGAGGCTTGTGAGCG 672
 731 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 790
 673 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732
 791 TTGATCTCTCAAGACTCCGCTTATGATGATGATGATGATGATGATGATGATGATGAT 850
 733 TTGATCTCTCAAGAGGCTGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 792
 851 TATCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 907
 793 TCGGCGCGCTCTCAAGACTCCGCTTATGATGATGATGATGATGATGATGATGATGAT 852
 908 TCGAGCTTGTGCT 967
 853 CCGTGAAGCTTCT 912
 968 TTCTTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1027
 913 TTCTCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 972
 1028 TCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
 973 TCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1032
 1088 TAGTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1147
 1033 TTGTTGGGCGGTGAGAGGTGGTGGCGAGCGCTGCGCAACTCTCTTCACTCCCAAG 1092
 1148 CTGTGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1207

1093 CCATGTCAGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1152
 1208 TGGCATTAAAGACAGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1267
 1153 TCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
 1268 TGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1327
 1213 TGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
 1328 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1387
 1273 CGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 1388 GACTTGAACCTTTCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAT 1447
 1321 GAGTGGCGCGTTCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAT 1380
 1448 CCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1507
 1381 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1437
 1508 ACGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1562
 1438 ATGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1492

RESULT 11

ABT23284

ID ABT23284 standard; DNA; 1545 BP.

XX ABT23284;

XX 01-MAY-2003 (first entry)

XX Seed development embryo/endosperm size alteration DNA SEQ ID 96.

XX Cytochrome P450; embryo; endosperm; size; seed development;

XX antisense inhibition; co-suppression; transformed plant; gene; ds.

XX Zea mays.

XX WO200239063-A2.

XX 12-DEC-2002.

XX 04-JUN-2002; 2002MO-US17562.

XX 05-JUN-2001; 2001US-295921P.

XX 28-NOV-2001; 2001US-334317P.

XX (DUPO) DU PONT DE NEMOURS & CO E. I.

XX Cahoon RE, Heppard EP, Nagaawa N, Sakai H;

XX WPI; 2003-229292/22.

XX P-PsDB; ABU26714.

XX New isolated nucleic acid fragment for antisense inhibition or

XX co-suppression of a cytochrome P450 polypeptide associated with

XX controlling embryo/endosperm size during seed development in plants

XX Claim 38; Page 127; 137pp; English.

XX The invention relates to an isolated nucleic acid fragment comprising a

XX promoter consisting essentially of a sequence having 2000, 2022, 8300 or

XX 17201 base pairs fully defined in the specification, or their similar and

XX functionally equivalent fragment or subfragment. The nucleic acid encodes

XX a cytochrome P450 polypeptide associated with controlling embryo/

XX endosperm size during seed development. The nucleic acid fragment is

XX useful in antisense inhibition or co-suppression of a cytochrome P450

XX polypeptide associated with controlling embryo/endosperm size during seed

XX development in a transformed plant. This polynucleotide sequence

CC represents a gene relating to the alteration of embryo/endosperm size
 CC during seed development of the invention.

XX Sequence 1545 BP, 182 A, 572 C, 559 G, 232 T, 0 other;

Query Match 18.0%; Score 286; DB 25; Length 1545;

Best Local Similarity 54.1%; Pred. No. 4,9e-86;
 Matches 692; Conservative 0; Mismatches 550; Indels 36; Gaps 4;

QY 306 CGAGCCAAAGAGCTTCATGCTTTTACGTTAGAGAGACTGCGTGAATGCTCAAGTCA 365
 DB CGGCGCCCGCGCTGTCGCGCTTCTCCGTCGCGCTCAAGCGCTCGCTGTAAGAGCA 314
 QY 366 TCCCGACGAGCGAAGAGATTCTGATAGCCCGGTTTGTGATCGACCGGTTAAGA 425
 DB GCCGACACAGCGCGCGGAGCTCTGCGACAGCGCGCTTGGCGACGCGCGCGGAGGA 374
 QY 426 ATCGCTTACTACTGATGTTTACAGAGCAATTGGTTTGCACACAGGTTTACTG 485
 DB CGCGCGCGCGGCGCTCTCTTCCACGCGCGCATGGGCTTTGCGCGCGCATCTAG 434
 QY 486 GCGAAGCTTGGCGCTATCGCTTGAACCATCTCTTTAGTACAAACAAATCAGAGAGC 545
 DB GCGCGCGCTTGGCGCGATCAGCTCGCGCTTCCGCTTCCAGCGCGCGCGAGCGCGG 494
 QY 546 CGAGACGACGACGAGTATCTCAAGCCAGATGAGTTTCTTGAAGAAACAGAGTAG 605
 DB GGGCGCGCGCGCGCGCGCATCGAGAGCGATGCGAGACCTCTCGCGCGCGCGG 554
 QY 606 TTAACGACCCCTTTTGTTCGAGTGTCTTAAACGCGCTGCTTAAACATGATAGT 665
 DB 555 ACGAAGAGTCC--GTATGCGCGCGCTGCTCCAGCGCGCATCTCGAGCACTGATGC 611
 QY 666 CTCTGATTCGAGCAAGATATAGCTTGAA--AAACATGTTGATTAAGTGAAT 722
 DB 612 CACCGTGTTCGCGCGCGCTTACAGCGCGCGCACCGCGAGCGCGGAGTGAAGAT 671
 QY 723 GGTCAAGAGGTTATGATTTGCTGGAACTTTGAATTGACTGATCACTTCTTGCT 782
 DB 672 GGTGAAGAGGTTAGACCTGCTGCGATGTTCAACTGGGCGACCACTGCGCTCT 731
 QY 783 ATCGAGTTTGTATCTCAAGACTCGGCTCTAGATTTCCACACTGTAACAAAGTAA 842
 DB 732 CAGGTGCTGAGACTGCAAGGCGTCAAGAGGCGGTGAGAGACCTGGTGGAGAGTCA 791
 QY 843 CCGGTTTGTATCCCGGATTAATCCGAACCCGTATCAACCGGTGATTTGCTCTGTA 902
 DB 792 CGTGTTCGTGCGCAGATATCGAAGAGCAGCGAAGAGCAGCGCATGAGGA 851
 QY 903 T-----TTCGACAGCTTTTGTCTCCCTCAATGTTTCAATTAAT 944
 DB 852 GCGCGCGCGCGCGGAGACTTCTGACGCTTCTGCGGATCGAGGCGGAGAGAGACT 911
 QY 945 ATTCGACCGGACATATGCGGCTTTTGGAGATGATTAATCAAGAGACAGACAGT 1004
 DB 912 GTGCGACTCGACATATGCTGTCTCTGAGAGATGATCTTTCAGGAGACGACAGCT 971
 QY 1005 TCGGCTTAATGAGTGAATCTGCTAGATGAGTCTTCAATCCAGATATCAATCAAC 1064
 DB 972 GGGGATCTCTGATGATGATGCGGATGCGGATGCGTCAACCGCGGATCAATCA 1031
 QY 1065 GTTACAAAGAGCTGATCAATGATCTGAGAAATCAAGACCTTATGATTAATCTGACT 1124
 DB 1032 GCGGACGCGGAGACTGACGCGCTGTGGGCGCGCGCGCGCTTTGCGAGCGGACGT 1091
 QY 1125 GGTTCATCTCATATCTAAGCGCTGTGGGAAAGAAAGTATGAGGCTTCACTCCAGG 1184
 DB 1092 GGGCGCGCTGCTTACTGACGCGCTGTGAAGAGAGAGCTCCGCTCAACCGCGG 1151
 QY 1185 CCCACTTTCATGATGGCGCGGTTGGCATTAACAGACAGATGTTGATGTCGTTGT 1244
 DB 1152 CCGGCTGCTCTGCTGGGCGCGCTGCGCGCGGACGCGGATGTCGCGGCGACAGTGT 1211

QY 1245 TCCGCGAGGAGCACACAGCAATGATGAGCGCGTATGCAATGATCAACGCTGTG 1304
 DB 1212 CCGCGGCGGACACAGCGCATGATGATGAGCGCATGCGGACGAGCCCGCGGTG 1271
 QY 1305 GGTGATCTTTGAGATTAACTGAGAGTTCGTGGCAAAAGAGTGAAGTGA 1364
 DB 1272 GCGGAGCGCTCCGCTTTCCGCGCGAGCGGTTCCAG-----GTGAGAGAGT 1319
 QY 1365 TTGCTTTTGGGATCGATTTGAGACTTGCACCTTCCGAGTGGGTCGCTGCGATTGCCC 1424
 DB 1320 GAGCGTCTGGCGCGGAGACTCGGCTCGCGCTTCCGCGCGCGCGCGCGGTGCCC 1379
 QY 1425 CCGGAAAGATCTTGTGTTTACTACCGTTATGTTTGAACGCGGATGATTAATGA 1484
 DB 1380 GGGCAAGACGCTGGCGCTGCGCATGTCACCTCTGCTGCGCAGCTGCGACCGCTT 1439
 QY 1485 TGAATGGGAGCGTCCGATGATTAACGCGGTTGATCTATCTGAGAACTGAGGCTTCTTG 1544
 DB 1440 CCGTGGCGCGCGCGGACGCGCGCGCGCTGCACTGCGGAGCGCGCTCGCATGCT 1499
 QY 1545 CGAGATGCTTAATCTCT 1562
 DB 1500 GAGATGAGAAAGCCCT 1517

RESULT 12

ABT23283 ID ABT23283 standard; DNA; 1758 BP.

AC ABT23283;

DT 01-MAY-2003 (first entry)

DE Seed development embryo/endosperm size alteration DNA SEQ ID 94.

KM Cytochrome P450; embryo; endosperm; size; seed development;

XX antisense inhibition; co-suppression; transformed plant; gene; ds.

OS Zea mays.

FN W0200299063-A2.

PD 12-DEC-2002.

PE 04-JUN-2002; 2002MO-US17562.

XX 05-JUN-2001; 2001US-295921P.

PR 28-NOV-2001; 2001US-334317P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

PI Caboon RE, Heppard BP, Nagasawa N, Sakai H;

DR WPI; 2003-228292/22.

XX P-PSDB; AB026713.

PT New isolated nucleic acid fragment for antisense inhibition or

XX co-suppression of a cytochrome P450 polypeptide associated with

XX controlling embryo/endosperm size during seed development in plants

XX Claim 38; Page 124-125; 137pp; English.

CC The invention relates to an isolated nucleic acid fragment comprising a

CC promoter consisting essentially of a sequence having 2000, 2022, 8300 or

CC 17201 base pairs fully defined in the specification, or their similar and

XX Sequence 1758 BP; 274 A; 650 C; 545 G; 289 T; 0 other;
 SQ

Query Match 17.8%; Score 283; DB 25; Length 1758;
 Best Local Similarity 55.1%; Pred. No. 5,6e-85;
 Matches 709; Conservative 0; Mismatches 530; Indels 48; Gaps 6;

QY 306 CGAGACCAAGAGCTCATGGCTTTCAGCTTAGAGAGATCCGGCATGTCAGCGTCAA 365
 DB 444 CGCGGCAAGGCACTATGACCTTCGGTGGGCTGACCCGCTGCTGCTGCTCAAGCA 503
 QY 366 TCCGACGTAAGCAAGATCTGATAGACCGGCTTTTGTGATGACCGGTTAAAGA 425
 DB 504 GCGCGATAGCGCGCGGAGATCTCGCACGCGCGGTTGCGGAGACCGCCGTCAGGA 563
 QY 426 ATGGCTTACTACTGATGTTTAAAGAGCAATGGTTTGCACCAAGG---TGTTTA 482
 DB 564 CGCGGCGCGCACTGCTCTTCCACGCGCGCATGGGCTTCGCGCTCGGAGACGCGCA 623
 QY 483 CTGCGCAAGCTTCGCGGCTTCGCTTGAACCATCTCTTTAGTCAAAAACAATCAGAAG 542
 DB 624 CTGGCGCGGCTCTCGCGGCTCTCGCGGCAACCACTGTTGCGCGCGCGCGCGGCGG 683
 QY 543 AGCCGAGACGCAACGAGATGATCTCAAGCCAGATGGTTGAGTTTCTGAAAAACAGAG 602
 DB 684 TGCCTGCGCACACCGCGCTCCATCGCGAGGCCATGGTGGCCGACGTCGCGCTGCCAT 743
 QY 603 TAGTAAAGAACCTGTTTGTGTGAGT---TGCTTAAAGCGGCTGCTTAAACAAT 659
 DB 744 GCGCGCGCACGCGAGGCTCCCTCTCAAGCGCTGCTGCAATGTCATCTCTCAACCACT 803
 QY 660 GATGTGCTCTGATTTGCAACAGATGATGAGCTTGAACCAACCATGTTGAGTTAGTGA 719
 DB 804 CATGGCCACCGTGTGGCAAGCGCTACACATGGGACGCGAGAGGCGCCCTTCTGGA 863
 QY 720 A---ATGCTGAAGAGATGATGATTTGCTCGGAAGTTGAATGATGATGATGATGATG 776
 DB 864 CGAGATGCTGCGGAGGCTTACGACCTCTGCGGCACTTCACTGAGCTGACCACTGCC 923
 QY 777 TTGGCTATCGGAGTTGATCTCTCAAGACTCCGCTAGATGTTTCCACACTCTGACCAA 836
 DB 924 ATTGCTCAAGCATCTGACCCCGAGGCGGCGGCGCGGCTGCAACAGGCTGCTCGAAA 983
 QY 837 GGTAAACCGGTTGTATCCGAGATTAATCCGAACACCGTAATCAACCGG----- 887
 DB 984 GGTGCAATCGTTCGTGGCAAGATCATCTTGGAGCACAGGCGCGCGGCGCAATGAGAG 1043
 QY 888 -----TGATTTGCTGCTGATTTGCTGAGCTTTTCTCTCCCTCCATGATTC 935
 DB 1044 AGTCGTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1103
 QY 936 AGATTAATTAATCCAGCCCGGACATATGCGCGCTTCTTGGAGATGATTAATCCAGAGAC 995
 DB 1104 GGAAGACCTGTCATATCGGACATGATGCTGTTCTTGGAGATGATGATGATGATGATG 1163
 QY 996 AGACACAGTTGCGGTTTAATCGAGTGAATCCTGCTAGGATGATGATGATGATGATGAT 1055
 DB 1164 CGAACCGTGGCATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1223
 QY 1056 GCATTAACCGGTACAAAACGAGCTGATCAATGATCGGGAATCAAGAGCCTTATGATGA 1115
 DB 1224 CCAAGCGCAAGCCCAAGCGGAGCTGAGCGGCTGAGGAGCTCGGCGCGCGCTGAGGGA 1283
 QY 1116 ATCTGACTTGGCTTCACTTCAATCAAGCGCTGCTGAGGAGATGATGATGATGATGATG 1175
 DB 1284 CGCGAGCTGCGCAAGCTTCACTTCAATCAAGCGCTGCTGAGGAGATGATGATGATGATG 1343
 QY 1176 TCTTCAGAGCCACTTCTATCATGAGCGCGCTTGGCCATTAACAGACAGATGTTGATGG 1235
 DB 1344 CCGGCAAGCGCGCTCTGCTGCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1403
 QY 1236 TCGCTTTTTCGCGAGGAGCAACAGATGATGATGATGATGATGATGATGATGATGATG 1295

DB 1404 CCACCTGTCGCCCGCGGACCAACAGCATGTAATGATGATGATGATGATGATGATGATG 1463
 QY 1296 ACAGCTGTGGTGTATCTTGTGAGTTAACTGAGAGGTTGTGGCAAAAGAGTGA 1355
 DB 1464 CGCCATGTGGGCGGAGCGAGAAATTCGCCCGCGAGCGGTTCCAG-----GA 1511
 QY 1356 GGTGAGATTTTCGTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1415
 DB 1512 GAGAGAGTGAAGCTCTCGGAGAGCACTCCGCTGAGCCCTTGGCGCGCGCGCGCG 1571
 QY 1416 GATTGCGCGCGGAGAAATCTTGTGTTTACTACCTTATGATGATGATGATGATGATG 1475
 DB 1572 CGCCTGCGCGGCAAGTACTGCGCTGCGCACACCACTCTGAGTGGCCCACTTCT 1631
 QY 1476 ACATGAGTTGAATGAGGAGCCGTCGATGTTAACGCGGCTGACTTATGAGAACTGAG 1535
 DB 1632 GCACAAATGTAAGTGG-----CGCGCGCGGCGGCGTGAAGAGCGCTGAG 1685
 QY 1536 GCTTCTTGGGAGATGCTATCTCT 1562
 DB 1686 CATGTCGCTGAGATGCGCACGCGCT 1712

RESULT 13

ID ABR23237 standard; DNA; 1597 BP.

ABR23237;
 01-MAY-2003 (first entry)

Seed development embryo/endosperm size alteration DNA SEQ ID 16.

Cytochrome P450; embryo; endosperm; size; seed development;
 antisense inhibition; co-suppression; transformed plant; gene; ds.

Zea mays.

WO200299063-A2.

12-DEC-2002.

04-JUN-2002; 2002MO-US17562.

05-JUN-2001; 2001US-295921P.

28-NOV-2001; 2001US-334317P.

(DUPO) DU PONT DE NEMOURS & CO E. I.

Cahoon RE, Hepard EP, Nagasawa N, Sakai H;

WPI: 2003-229292/22.

P-FSDB; ABR26681.

New isolated nucleic acid fragment for antisense inhibition or

controlling embryo/endosperm size during seed development in plants

Claim 38; Page 83; 137p; English.

The invention relates to an isolated nucleic acid fragment comprising a promoter consisting essentially of a sequence having 2000, 2022, 8300 or 17201 base pairs fully defined in the specification, or their similar and functionally equivalent fragment or subfragment. The nucleic acid encodes a cytochrome P450 polypeptide associated with controlling embryo/ endosperm size during seed development. The nucleic acid fragment is useful in antisense inhibition or co-suppression of a cytochrome P450 polypeptide associated with controlling embryo/endosperm size during seed development in a transformed plant. This polynucleotide sequence represents a gene relating to the alteration of embryo/endosperm size during seed development of the invention.

Sequence 1597 BP; 279 A; 523 C; 517 G; 278 T; 0 other;

Query Match 17.5%; Score 278.2; DB 25; Length 1597;
 Best Local Similarity 54.9%; Pred. No. 2,3e-83;
 Matches 706; Conservative 0; Mismatches 533; Indels 48; Gaps 6;

QY 306 CGAGAGCAAGAGCTCATGCTTTCAGCTTAAAGAGAGCTGCGTGAATGCTCACTGCA 365
 DB CGAGGAGCAAGAGCTCATGCTTTCAGCTTAAAGAGAGCTGCGTGAATGCTCACTGCA 109
 QY 366 TCCGAGCTGAGAGAGAGATTCGATAGCCCGGTTTGTGATGCAACCGTTTAAAG 425
 DB CGCGGAGCAAGAGCTCATGCTTTCAGCTTAAAGAGAGCTGCGTGAATGCTCACTGCA 169
 QY 110 GCGGAGCAAGAGCTCATGCTTTCAGCTTAAAGAGAGCTGCGTGAATGCTCACTGCA 169
 DB 426 ATCGGCTTACTGCTGATGTTTAAAGAGAGATGTTTGCACACAGG--TGTTTA 482
 QY 170 CGCGGAGCAAGAGCTCATGCTTTCAGCTTAAAGAGAGCTGCGTGAATGCTCACTGCA 229
 DB 483 CTGCGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTGCGTGAATGCTCACTGCA 542
 QY 230 CTGCGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTGCGTGAATGCTCACTGCA 289
 DB 543 AGCGGAGCAAGAGCTCATGCTTTCAGCTTAAAGAGAGCTGCGTGAATGCTCACTGCA 602
 QY 290 TCGCGGAGCAAGAGCTCATGCTTTCAGCTTAAAGAGAGCTGCGTGAATGCTCACTGCA 349
 DB 603 TAGTAAAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTGCGTGAATGCTCACTGCA 659
 QY 350 GCGCGGAGCAAGAGCTCATGCTTTCAGCTTAAAGAGAGCTGCGTGAATGCTCACTGCA 409
 DB 660 GATGCTCTGATTCGAGAGAGATGAGCTTGAAGAGAGCTTGAAGAGAGCTTGAAGAG 717
 QY 410 CATGCGGAGCAAGAGCTCATGCTTTCAGCTTAAAGAGAGCTGCGTGAATGCTCACTGCA 469
 DB 718 -GAAATGCTGAGAGAGATGAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAG 776
 QY 470 CGAGAGCTGAGAGAGCTCATGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAG 529
 DB 777 TTGCTATCGAGAGAGCTCATGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAG 836
 QY 530 ATTGCTCAAGAGCTCATGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAG 589
 DB 837 GGTAAAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAG 887
 QY 590 GGTAAAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAG 649
 DB 888 -----TGATTTGCTGATTTGCTGAGAGCTTTCAGCTTTCAGCTTTCAGAGTTTC 935
 QY 650 ACTGCTGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAG 709
 DB 936 AGATTAATTCAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAG 995
 QY 710 GAGAGAGCTGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAG 769
 DB 996 AAGACAGCTGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAG 1055
 QY 770 CGAGAGCTGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAG 829
 DB 1056 GATTCAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAG 1115
 QY 830 CGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAGAGCTTTCAG 889
 DB 1116 ATCTGCTTTCAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAGAGCTTTCAG 1175
 QY 890 CGCGAGCTGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAGAGCTTTCAG 949
 DB 1176 TCTTCAGAGAGCTTTCAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTTCAG 1235
 QY 950 CCGCGAGAGAGCTTTCAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTTCAG 1009
 DB 1236 TCGTCTTTCAGAGAGCTTTCAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTTCAG 1295
 QY 1010 CCACTGCTTCAGAGAGCTTTCAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTTCAG 1069

QY 1296 ACAGCTGCTGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTGAAGAG 1355
 DB 1070 CGCGAGCTGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTTCAG 1117
 QY 1356 GATGAGAGCTTTCAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTTCAG 1415
 DB 1118 GAGAGAGCTGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTTCAG 1177
 QY 1416 GATTTGCTGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTTCAG 1475
 DB 1178 CGCGAGCTGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTTCAG 1237
 QY 1476 ACATGAGCTTTCAGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTTCAG 1535
 DB 1238 GAGAGAGCTTTCAGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTTCAG 1291
 QY 1536 GCTTCTTCAGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTTCAG 1562
 DB 1292 CATGCTGCTGAGAGAGCTTTCAGCTTTCAGCTTAAAGAGAGCTTGAAGAGAGCTTTCAG 1318

RESULT 14
 ABT23282
 ID ABT23282 standard; DNA; 1585 BP.
 XX
 AC ABT23282;
 XX
 AC 01-MAY-2003 (first entry)
 DT
 XX
 DE Seed development embryo/endosperm size alteration DNA SEQ ID 92.
 XX
 KM Cytochrome P450; embryo; endosperm; size; seed development;
 XX antisense inhibition; co-suppression; transformed plant; gene; ds.
 XX
 OS Hordeum vulgare.
 XX
 PN WO200299063-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 04-JUN-2002; 2002WO-US17562.
 XX
 PR 05-JUN-2001; 2001US-295921P.
 XX
 PR 28-NOV-2001; 2001US-334317P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Cahoon RE, Heppard BP, Nagasawa N, Sakai H;
 XX
 DR WPI; 2003-229292/22.
 XX
 DR P-PSDB; AB26712.
 XX
 PT New isolated nucleic acid fragment for antisense inhibition or
 PT co-suppression of a cytochrome P450 polypeptide associated with
 PT controlling embryo/endosperm size during seed development in plants
 XX
 PS Claim 2; Page 122; 137pp; English.
 XX
 XX The invention relates to an isolated nucleic acid fragment comprising a
 CC promoter consisting essentially of a sequence having 2000, 2022, 8300 or
 CC 17201 base pairs fully defined in the specification, or their similar and
 CC functionally equivalent fragment or subfragment. The nucleic acid encodes
 CC a cytochrome P450 polypeptide associated with controlling embryo/
 CC endosperm size during seed development. The nucleic acid fragment is
 CC useful in antisense inhibition or co-suppression of a cytochrome P450
 CC polypeptide associated with controlling embryo/endosperm size during seed
 CC development in a transformed plant. This polynucleotide sequence
 CC represents a gene relating to the alteration of embryo/endosperm size
 CC during seed development of the invention.
 XX
 XX Sequence 1585 BP; 228 A; 551 C; 523 G; 283 T; 0 other;
 XX
 Query Match 17.3%; Score 275.2; DB 25; Length 1585;

Best Local Similarity 54.7%; Pred. No. 2,4e-82;
Matches 706; Conservative 0; Mismatches 533; Indels 51; Gaps 6;

QY 306 CGAGCCGCAAGAGCTCTGATGCTTCACTTGAAGAGATCGCTGATGTCAGTCGCA 365
Db CGAGCCGCGGCGATGATGCTCTTCCGCTCGGCTCAACCCGCTGCTGCTGAGCAG 348
QY 366 TCCGACGTAAGGAAAGATCTGATATGCCCCGGTTTGTGATTCAGCCGTTAAAG 425
Db GCAAGATACGCGCGCGAATCTGCTCAACCCGCGTTCCGCGACCGCCGCTGAAGA 408
QY 426 ATCGCTTACTCATGATGTTTAAACAGACAAATGGTTTGCAC---CACAGCTGTTA 482
Db CCGGGGCGCCACCTCTCTTCCACCGCCGCAATGGGTTTGCCTGCGGCGAGCGCCA 468
QY 483 CTGCGCAAGCTTCCGCGCTATCGCTTCGAACATCTCTTATGACAAACAAATGAGA 542
Db CTGGCGTGGCTGGCGCTGCTCGCGCGCACCTCTCGGCTCTCGCGCTGCGCGGCT 528
QY 543 AGCCGAGACGCAACGAGTGTAT---CTCAAGCCAGATGTTGATTTCTTGAACAA 599
Db CTCCGACACCCACCTCTCTATTTGGGCGGCGCATGCTCGGCGAGCTGCTCATCAT 588
QY 600 GAGTATTAACGAACCTGTTTGTGATGTTGCTTAAACGGCGTCCGTTAAACAT 659
Db GGGCCCGCACGCGGAGTGCCTCTCGAGGTTCTGACAGCGGCTCTCTCAACACAT 648
QY 660 GATGCTCTGATTCGACCAAGATGATGCTTGAACAAACCA-----TGTGATTT 713
Db CATGCGCTGCTTTTGGCAAGCGCTACACACCTTCAACAGCAAGAGAGTGTGT 708
QY 714 ACCTGAATGCTCGAAAGATTAATTTGCTCGAAGCTTGAATTTGATGATGATCCT 773
Db GGAGGAGATGCTTAAACGAAGGATACACCTCTCGGCAAGTTCACCTGGGAGATCACCT 768
QY 774 TCCCTGGCTATCGGAATTTGATCTCAAGACCTCGGATGATGTTCCACCTGTACC 833
Db GCCATTCTTAAGTCTGCTCATCTCAAGGCGCTGCGGCGCGGCTGCAACAGTTAGTCCG 828
QY 834 AAAGTAAACCGGTTTGTATCCCGATTAATATCCGAACACCTGATCAACACCGG----- 887
Db GCAAGTGAAGGCTGATCGTCAATCAATACAGAGACCAAGGCGAGCGACAGTGC 888
QY 888 -----TGATTTGCTGCTGATTTGCTGCGAGTTCCTGCTCTCCCTCATG 932
Db ATAGGCTATTCGAGATGATCTCTCGGCGACTTCTGCTGATGCTCTCTCGGCTGACGG 948
QY 933 TTCGATTAATATTCGACCCCGACATTAATCGCGCTTTTGGAGATGATATTTCAGCG 992
Db AGAAGACAAATGTCAGAGTCCGACATGATCGCGCTTTTGGAGATGATCTTTAGAGG 1008
QY 993 AACAGACAGTTCGCTTTTATTCAGATGATCTCTGCTAGATGCTCTTATATCAAA 1052
Db GACGGAACAGTGGCTGATGAGATTAATGCGAGATGATGCTGACCCGGA 1068
QY 1053 TATGATCAACGCTGACAAACAGCTGATCAAGTATGCGGAAATCAAGACCTTAAG 1112
Db GATTCATGAAAGCCCGGCGGAGAGCTTGAACGCGTGTGGCGGCGGCGCTGAGC 1128
QY 1113 TGAATCTGACTGGCTTCACTTCCATTTCAAGGCTGTGTGAAGAAATATTTAGGCT 1172
Db GAGAGAGAGCTGTGAGGCTCTCCATCAATCCAGTGTGTAAGAGACGCTGGCAT 1188
QY 1173 TCATCTCCAGGCGCACTTCTATCATGCGCCGTTTGGCATTAACAGACATGCTTGA 1232
Db GCACCCCGCGGCGCTCTCTCTATGAGGCGGCTGCGTGAACAGACGCGACCTTCG 1248
QY 1233 TGGTGTCTTGTTCGCGAGGACACAGCATGATGATGATGATGATGATGATGAT 1292
Db CGGCACTGCTGCTGCGCGCGACAGCGGATGATGATGATGATGATGATGATGAT 1308
QY 1293 TCCACAGTGTGCTGATCTTGTGAGTTAACTGAGAGGTTGCTGCAAAAGAGG 1352

Db 1309 CGCGCGGTGTGGCCGAGCCGAGCTGTTCCGCGCGAGCGGTTGATG----- 1357
QY 1353 TGAGTGAATTTTCGTTCTTGGGTCGCACTTTGACACTTTCGAGTGGGTCGGGTCG 1412
Db 1358 -GAGAGAGAGTGAAGCTGTGGGACAGCACTTCCGCTTGCCCCGTTCCGCGCCGCG 1416
QY 1413 TCGGATTTGCCCGGGAAGATCTTGTATTTACTACCGTTATGTTTGAACGCGCATGAT 1472
Db 1417 GCGGTGTGCCCGGGAAGATGCTGGCCCTGCGCACCGTCACTGCTCGCGACGCT 1476
QY 1473 GTTACATGATTTGAATGGGACACCTCCGATGTTAAGCGGCTGATCTTATCTGAGAACT 1532
Db 1477 GCTTACACCGGTGTGAGTGGGCTCCCTCGG-----GAGCGTGAACCTGTCAAGAGGCT 1530
QY 1533 GAGGCTTCTTGCAGATGAGCTATCTCT 1562
Db 1531 CAAGATGTCACGTGAATGCGCACGCGCT 1560

RESULT 15
ABT23238
ID ABR23238 standard; DNA; 1539 BP.

AC ABR23238;
DT 01-MAY-2003 (first entry)

DE Seed development embryo/endosperm size alteration DNA SEQ ID 18.

XX Cytochrome P450; embryo; endosperm; size; seed development;
KW antisense inhibition; co-suppression; transformed plant; gene; ds.

XX Zea mays.

XX MO20029063-A2.

XX 12-DEC-2002.

XX 04-JUN-2002; 2002MO-US17562.

XX 05-JUN-2001; 2001US-295921P.

XX 28-NOV-2001; 2001US-334317P.

XX (DUPO) DU PONT DE NEMOURS & CO E. I.

PI Canoon RE, Heppard EP, Nagasawa N, Sakai H;

XX WPI; 2003-229292/22.

DR P-PSDB; ABJ26682.

XX New isolated nucleic acid fragment for antisense inhibition or

PT co-suppression of a cytochrome P450 polypeptide associated with

XX controlling embryo/endosperm size during seed development in plants

XX Claim 38; Page 85; 137pp; English.

XX The invention relates to an isolated nucleic acid fragment comprising a

CC promoter consisting essentially of a sequence having 2000, 2022, 8300 or

CC 17201 base pairs fully defined in the specification, or their similar

CC functionally equivalent fragment or subfragment. The nucleic acid encodes

CC a cytochrome P450 polypeptide associated with controlling embryo/

CC endosperm size during seed development. The nucleic acid fragment is

CC useful in antisense inhibition or co-suppression of a cytochrome P450

CC polypeptide associated with controlling embryo/endosperm size during seed

CC development in a transformed plant. This polynucleotide sequence

CC represents a gene relating to the alteration of embryo/endosperm size

CC during seed development of the invention.

XX Sequence 1539 BP; 274 A; 481 C; 485 G; 298 T; 1 other;

XX Query Match 17.2%; Score 274.4; DB 25; Length 1539;

XX Best Local Similarity 60.3%; Pred. No. 4.4e-82;

XX Matches 524; Conservative 0; Mismatches 312; Indels 33; Gaps 3;

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rng

Page 18

Db	1078	AAGCTGTCTGTGGAGATGGCCGTCCGCT	1106
Qy	1534	AGGCTTTCTTGGAGATGGCTATCTCT	1562
Db	1027	CTGCACGAGTTTGAAGTGTCTCCCTC-----GCCGTGCACCTGTGGAGAGTGTCTC	1077
Qy	1474	TTACATGAGTTTGAATGGGACACCGTCCAGTGTAAACGGGGTGTGACTTATCTGAGAACTG	1533
Db	967	CGACACTGCCCGGAGAGAGCTCCACATGAGCCACCGTGGCCCTTCTGTCTGCCGACGCTG	1026
Qy	1414	CGAATTTGCCCGGAGAGAAATCTGTGTTTACTACCGTTATGTTTGAACGGCGATGATG	1473
Db	907	TCCACCGAGTTTCCCATCATGAGAGTCCGAACTCAGAGCTCGGCCGTTTGGGGCGGGCCGG	966
Qy	1354	GAGGTGAGAGTTTCCGTTCTTGGGTGGATTTGACATTTGACATTTGACACTTGGGGTCGGGTGCT	1413
Db	856	CCAGACGTGTGGGCGAGCCCGACGAGATTCCAGCGCGAGAGTTTCATG-----GGG	906
Qy	1294	CCACACGTGTGGGTGATCTTTTGGAGTTTAAACCTGAGAGTTCTGTGCAAAAGAGGT	1353
Db	796	GAGCACTCATCTCCCGCGGAGACACCGGATGTGTGAACATGTGGGCCATTATCCGACAC	855
Qy	1174	CATCTCCAGGCCCACTTCTATCATGTGGGCCCGTTTGGCCATTAACAGACGATCGTGTAT	1233
Db	736	CACCGCGCGGCGCTGTCTGTGCGGGCGCGCTTGCGCACGTCCAGATGATCACTGCAC	795
Qy	1224	GATCGTCTTTTCCCGGAGAGGACCAACACATATGTGAAACATGTGGGCCGTATTCAGTAT	1293
Db	796	GAGCACTCATCTCCCGCGGAGACACCGGATGTGTGAACATGTGGGCCATTATCCGACAC	855
Qy	1054	ATGCATATCAGAGTACAAACGAGGTGATCAAGTAGTCGGGAAATCAGAGCCCTGAT	1113
Db	616	GTGCAGGCCCGCTGCACGAGAGCTGAGACCGCGGTGCGGCGCCGACCGGCGGTCAAC	675
Qy	1114	GAATCTGACTTGGCTTCACTTCCATATTAACGCTGTGTGGAAAGATTTAGGCTT	1173
Db	676	GAGTCCGACACGGGCTCACTGTCTACTGTGCACGCGGTATCAGAGAGTCTCTCAGATG	735
Qy	994	ACAGACACAGTTGGGCTTATATGAGTGAATCTCTGTAGAGATGTCCTTCAATCCAGAT	1053
Db	556	ACTGACACCGTCCGCTGTCTATGAGTGGGCGCTGGCGACGCTGTGTGTGCACCTGTAC	615
Qy	1054	ATGCATATCAGAGTACAAACGAGGTGATCAAGTAGTCGGGAAATCAGAGCCCTGAT	1113
Db	616	GTGCAGGCCCGCTGCACGAGAGCTGAGACCGCGGTGCGGCGCCGACCGGCGGTCAAC	675
Qy	934	TCAGATAAATATCCGACCCGAGACATATCCCGTTCTTTTGGAGATGATATTCAGAGA	993
Db	496	GACGACAGGCTCACCGACGCTGACATGATGCCCGTCCGTCTGGGAAATGTGTTCGTGGA	555
Qy	889	GATTTGCTCGT-----GATTTGCTGCAGCTTTTGTCTCTCTCCCTCATGT	933
Db	436	CTCCGCTCTGTCCGCGCTGCGCGGTGAGACTTCCACACGCTCTGCTCTCGCTGCCGCC	495
Qy	316	CACCTCCCGCGCTCGCGTGTCTGACCTGACCTGACGACACAGGGCAGATGTGACCCGCTC	375
Db	829	GTACCAAAAGTAAACCGGTTTGTATCCCGAGTTATATCCGACACACCGTAAATCAAAACCGT	888
Qy	376	GTCCCGCTCGTGAACCGCTTGTGTGGCGGACATGTGCACAGACACCGCGCCGGAACAC	435
Db	889	GATTTGCTCGT-----GATTTGCTGCAGCTTTTGTCTCTCTCCCTCATGT	933
Qy	934	TCAGATAAATATCCGACCCGAGACATATCCCGTTCTTTTGGAGATGATATTCAGAGA	993
Db	496	GACGACAGGCTCACCGACGCTGACATGATGCCCGTCCGTCTGGGAAATGTGTTCGTGGA	555
Qy	994	ACAGACACAGTTGGGCTTATATGAGTGAATCTCTGTAGAGATGTCCTTCAATCCAGAT	1053
Db	556	ACTGACACCGTCCGCTGTCTATGAGTGGGCGCTGGCGACGCTGTGTGTGCACCTGTAC	615
Qy	1054	ATGCATATCAGAGTACAAACGAGGTGATCAAGTAGTCGGGAAATCAGAGCCCTGAT	1113
Db	616	GTGCAGGCCCGCTGCACGAGAGCTGAGACCGCGGTGCGGCGCCGACCGGCGGTCAAC	675
Qy	1114	GAATCTGACTTGGCTTCACTTCCATATTAACGCTGTGTGGAAAGATTTAGGCTT	1173
Db	676	GAGTCCGACACGGGCTCACTGTCTACTGTGCACGCGGTATCAGAGAGTCTCTCAGATG	735
Qy	1174	CATCTCCAGGCCCACTTCTATCATGTGGGCCCGTTTGGCCATTAACAGACGATCGTGTAT	1233
Db	736	CACCGCGCGGCGCTGTCTGTGCGGGCGCGCTTGCGCACGTCCAGATGATCACTGCAC	795
Qy	1224	GATCGTCTTTTCCCGGAGAGGACCAACACATATGTGAAACATGTGGGCCGTATTCAGTAT	1293
Db	796	GAGCACTCATCTCCCGCGGAGACACCGGATGTGTGAACATGTGGGCCATTATCCGACAC	855
Qy	1294	CCACACGTGTGGGTGATCTTTTGGAGTTTAAACCTGAGAGTTCTGTGCAAAAGAGGT	1353
Db	856	CCAGACGTGTGGGCGAGCCCGACGAGATTCCAGCGCGAGAGTTTCATG-----GGG	906
Qy	1354	GAGGTGAGAGTTTCCGTTCTTGGGTGGATTTGACATTTGACATTTGACACTTGGGGTCGGGTGCT	1413
Db	907	TCCACCGAGTTTCCCATCATGAGAGTCCGAACTCAGAGCTCGGCCGTTTGGGGCGGGCCGG	966
Qy	1414	CGAATTTGCCCGGAGAGAAATCTGTGTTTACTACCGTTATGTTTGAACGGCGATGATG	1473
Db	967	CGACACTGCCCGGAGAGAGCTCCACATGAGCCACCGTGGCCCTTCTGTCTGCCGACGCTG	1026
Qy	1474	TTACATGAGTTTGAATGGGACACCGTCCAGTGTAAACGGGGTGTGACTTATCTGAGAACTG	1533
Db	1027	CTGCACGAGTTTGAAGTGTCTCCCTC-----GCCGTGCACCTGTGGAGAGTGTCTC	1077
Qy	1534	AGGCTTTCTTGGAGATGGCTATCTCT	1562
Db	1078	AAGCTGTCTGTGGAGATGGCCGTCCGCT	1106

Search completed: January 17, 2004, 04:23:05
Job time : 456 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 15:28:56 ; Search time 4298 Seconds
(without alignments)
11360.561 Million cell updates/sec

Title: US-10-022-025A-1
Perfect score: 2009

Sequence: 1 aataataataataaatgta.....atagagaatgagatttc 2009

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 30

Total number of hits satisfying chosen parameters: 9

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_huv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	168	8.4	168	29	CC455276 SALK 0800
C 2	79	3.9	168	9	AV807688 AV807688
C 3	50	2.5	252	29	BZ380060 BZ380060
C 4	43	2.1	674	28	BH451907 BH451907

ALIGNMENTS

C	5	43	2.1	745	28	BH510149
C	6	38	1.9	82	29	AL946977 Arabidops
C	7	37	1.8	809	29	BZ428529 BONHW74TR
C	8	35	1.7	712	29	BZ428523 BONHW74TR
C	9	32	1.6	804	28	BH431683 BOHRT46TR

RESULT 1
CC455276/c
LOCUS
DEFINITION SALK 080087.38.10.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_080087.38.10.x, genomic survey sequence.

ACCESSION CC455276
VERSION CC455276
KEYWORDS GI:31215237

ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana
Bukariyoti Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab , C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.

TITLE A Sequence-indexed Library of Insertion Mutations in the Arabidopsis Genome

JOURNAL Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379

COMMENT Email: eckersalk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated intron of At2g46660.
Class: TDNA tagged.

FEATURES
location/Qualifiers
1..168
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK 080087.38.10.x"
/clone_id="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/cdna_protocols.html

BASE COUNT 70 a 25 c 11 g 62 t

ORIGIN
Query Match 8.4%; Score 168; DB 29; Length 168;
Best Local Similarity 100.0%; Pred. No. 2.4e-73;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1158 TAATTAATCTTAATTAATTTGGAATTCCTAGATTAATTAATTAACATGAA 1217
DB 168 TAATTAATCTTAATTAATTTGGAATTCCTAGATTAATTAATTAACATGAA 109

QY 1218 ATTGAAGCTCAATGTAAGTGTGGAATTAATTAAGCATTAAGATTGTTTGA 1277
DB 108 ATTGAAGCTCAATGTAAGTGTGGAATTAATTAAGCATTAAGATTGTTTGA 49

QY 1278 CATCAATTAATTAATTTGATTTTAAACAGTTTCTATTAAGTAAGTA 1325
DB 48 CATCAATTAATTAATTTGATTTTAAACAGTTTCTATTAAGTAAGTA 1

RESULT 2
AV807688/c 436 bp mRNA linear EST 29-MAR-2002
LOCUS AV807688
DEFINITION Arabidopsis thaliana cDNA clone RAFL09-51-118 3',
mRNA sequence.
ACCESSION AV807688
VERSION AV807688.1 GI:19841673
KEYWORDS EST,
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 436)
REFERENCE
AUTHORS Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Aizawa, T., Shibata, K., Shingawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished
TITLE JOURNAL
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekic@r.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FIC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified Bluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.

FEATURES
source
1..436
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL09-51-118"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_11b="RAFL9"
/note="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

BASE COUNT 168 a 92 c 56 g 120 t
ORIGIN

Query Match 3.9%; Score 79; DB 9; Length 436;
Best Local Similarity 100.0%; Pred. No. 2,7e-28;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1931 GATGGCTATCTCTTCCTGCTTAATGCGCGTAGGCGCAAGTAAAAAGAGCTCA 1930
DB 319 GATGGCTATCTCTTCCTGCTTAATGCGCGTAGGCGCAAGTAAAAAGAGCTCA 260
QY 1991 TATGAGAAATTAGAGATTTT 2009
DB 259 TATGAGAAATTAGAGATTTT 241

RESULT 3
BZ380060 262 bp DNA linear GSS 26-NOV-2002
LOCUS BZ380060
DEFINITION SALX_114536.29.00.x Arabidopsis thaliana TDNA insertion lines
survey sequence.
ACCESSION BZ380060
VERSION BZ380060.1 GI:25472537

KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 262)
REFERENCE
AUTHORS Alonso, J.M., Leisner, T.J., Barajas, P., Chen, H., Cheuk, R., Gadgil, N.,
C. Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P.,
Zimmerman, J. and Ecker, J.R.
A Sequence-indexed library of Insertion Mutations in the
Arabidopsis Genome
Unpublished
CONTACT: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated intron of At2g46660.
Class: TDNA tagged.

FEATURES
source
1..262
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALX_114536.29.00.x"
/clone_11b="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/cdna_protocols.html

BASE COUNT 64 a 56 c 62 g 75 t 5 others
ORIGIN

Query Match 2.5%; Score 50; DB 29; Length 262;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1495 ATGAATCTGACTTGGCTTCACTTCATATCTTAAGCGCTGTGTGAAGAA 1544
DB 160 ATGAATCTGACTTGGCTTCACTTCATATCTTAAGCGCTGTGTGAAGAA 209

RESULT 4
BH451907/c 674 bp DNA linear GSS 12-DEC-2001
LOCUS BH451907
DEFINITION BOGS044TR BOGS Brassica oleracea genomic clone BOGS044, genomic
survey sequence.
ACCESSION BH451907
VERSION BH451907.1 GI:17637618
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 674)
REFERENCE
AUTHORS Town, C.D., Van Aken, S., Uterback, T., Koo, H. and Frazer, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
OTHER GSS: BOGS044TF
CONTACT: Chris Town
TIGR Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

FEATURES
source

1..674

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/strain="TO1000DH3"

/db_xref="taxon:3712"

/clone="BOG504"

/note="Vector: PHOS1, Site 1: BstXI, 2-3 kb sheared

/note="genomic DNA inserted into PHOS1 using BstXI linkers"

191 a 194 c 130 g 159 t

BASE COUNT

191 a 194 c 130 g 159 t

ORIGIN

Query Match 2.1%; Score 43; DB 28; Length 674;

Best Local Similarity 100.0%; Pred. No. 4.6e-10;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1669 CGCATGATCCACACGTGTGGTTCATTCCTTGGAGTTAAACC 1711

DB 410 CGCATGATCCACACGTGTGGTTCATTCCTTGGAGTTAAACC 368

RESULT 5

LOCUS

BH510149/c 745 bp DNA linear GSS 13-DEC-2001

DEFINITION BOGT049R BOGT Brassica oleracea genomic clone BOGT049, genomic

survey sequence.

ACCESSION BH510149

VERSION BH510149.1 GI:11718239

KEYWORDS

SOURCE GSS.

ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eustroids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 745)

Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished

Other GSSs: BOGT049TF

COMMENT Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..745

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/strain="TO1000DH3"

/db_xref="taxon:3712"

/clone="BOGT049"

/clone_1lb="BOGT"

/note="Vector: PHOS1, Site 1: BstXI, 2-3 kb sheared

/note="genomic DNA inserted into PHOS1 using BstXI linkers"

219 a 186 c 137 g 203 t

BASE COUNT

219 a 186 c 137 g 203 t

ORIGIN

Query Match 2.1%; Score 43; DB 28; Length 745;

Best Local Similarity 100.0%; Pred. No. 4.6e-10;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1669 CGCATGATCCACACGTGTGGTTCATTCCTTGGAGTTAAACC 1711

DB 197 CGCATGATCCACACGTGTGGTTCATTCCTTGGAGTTAAACC 155

RESULT 6

LOCUS

AL946977 82 bp DNA linear GSS 24-OCT-2002

DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-300D11-015600,

genomic survey sequence.

ACCESSION AL946977

VERSION AL946977.1 GI:24403599

KEYWORDS

SOURCE GSS.

ORGANISM Arabidopsis thaliana (thale cress)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eustroids II; Brassicales; Brassicaceae; Arabidopsids.

1 Strizhov, N., Li, Y., Rosso, M., Vahoevar, P., Dekker, K., Saedler, H.

and Weisshaar, B.

A pipeline for automated high-throughput generation of FSTs

(flanking sequence tags) from Arabidopsis thaliana T-DNA

transformed lines

Unpublished

2

Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weisshaar, B.

A new Arabidopsis thaliana T-DNA mutagenesis population (GABI-Kat)

for flanking sequence tag based reverse genetics

Unpublished

3 (bases 1 to 82)

Li, Y., Strizhov, N., Rosso, M. and Weisshaar, B.

Direct Submission

Submitted (21-OCT-2002) Weisshaar B., Max-Planck-Institut fuer

Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany

This sequence is recovered from the left border of the T-DNA. It

indicates an insertion close to or within gene At2g46660. The

sequences are generated at the MPI for Plant Breeding Research in

the context of the GABI-Kat project. GABI-Kat is part of the German

Plant Genomics program designated 'GABI'. Information on line

availability can be found at:

http://www.mpiz-koeln.mpg.de/GABI-Kat/.

Location/Qualifiers

1..82

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="GK-300D11-015600"

/clone_1lb="Arabidopsis thaliana T-DNA insertion lines"

/note="PCR was performed on DNA from Arabidopsis thaliana

plants (Ti) which were transformed with the T-DNA from

vector pACT61. The lines contain one or more T-DNA

insertions. The DNA fragment(s) resulting from the PCR

were directly sequenced to determine the genomic sequence

flanking the insertion. Sequences displaying significant

similarity to the A. thaliana nuclear genome sequence were

processed for submission. T-DNA derived sequences were

removed"

BASE COUNT 17 a 13 c 12 g 19 t 21 others

ORIGIN

Query Match

Best Local Similarity 1.9%; Score 38; DB 29; Length 82;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1379 AGACACAGTTCGGTCTTAATCGAGTGATCTCGCTA 1416

DB 45 AGACACAGTTCGGTCTTAATCGAGTGATCTCGCTA 82

RESULT 7

LOCUS

B2428529/c 809 bp DNA linear GSS 13-DEC-2002

DEFINITION BONHW74TR BO.1.6.2 KB. tot Brassica oleracea genomic clone BONHW74,

genomic survey sequence.

ACCESSION B2428529

B2428529

BONHW74TR BO.1.6.2 KB. tot Brassica oleracea genomic clone BONHW74,

genomic survey sequence.

ACCESSION B2428529

B2428529

VERSION B2428523.1 GI:26671510
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
AUTHORS ; eurosids II; Brassicales; Brassicaceae; Brassica.
JOURNAL 1 (bases 1 to 809)
COMMENT Town,C.D., Van Aken,S., Uteback,T., Koo,H. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other_GSSs: BONHW74TF
Contact: Chris Town

TIGR
712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
source
1. 809
Location/Qualifiers

/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BONHW74"
/note="Vector: PHOS1; Site 1: BstXI, 1.6-2 kb sheared
total DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 241 a 216 c 146 g 206 t
ORIGIN

Query Match 1.8%; Score 37; DB 29; Length 809;
Best Local Similarity 100.0%; Pred.No. 5e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1627 TTCCGGCAGGAGCAGCAGCATGATGATGAGC 1663
DB 473 TTCCGGCAGGAGCAGCAGCATGATGATGAGC 437

RESULT 8
B2428523 712 bp DNA linear GSS 13-DEC-2002
LOCUS BONHW74TF BO.1.6.2 KB tot Brassica oleracea genomic clone BONHW74,
DEFINITION genomic survey sequence.
ACCESSION B2428523
VERSION B2428523.1 GI:26671493
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
AUTHORS ; eurosids II; Brassicales; Brassicaceae; Brassica.
JOURNAL 1 (bases 1 to 712)
COMMENT Town,C.D., Van Aken,S., Uteback,T., Koo,H. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other_GSSs: BONHW74TF
Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TP
Class: sheared ends.
FEATURES
source
1. 712
Location/Qualifiers

/organism="Brassica oleracea"

/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BONHW74"
/note="Vector: PHOS1; Site 1: BstXI, 1.6-2 kb sheared
total DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 168 a 179 c 177 g 188 t
ORIGIN

Query Match 1.7%; Score 35; DB 29; Length 712;
Best Local Similarity 100.0%; Pred.No. 5.2e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 797 AAACGGCGTCGCTTACACATGATGCTGT 831
DB 387 AAACGGCGTCGCTTACACATGATGCTGT 421

RESULT 9
B2431683/c 804 bp DNA linear GSS 12-DEC-2001
LOCUS BOHR746TR BOHR Brassica oleracea genomic clone BOHR746, genomic
DEFINITION survey sequence.
ACCESSION B2431683
VERSION B2431683
KEYWORDS B2431683.1 GI:17617404
SOURCE GSS.
ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
AUTHORS ; eurosids II; Brassicales; Brassicaceae; Brassica.
JOURNAL 1 (bases 1 to 804)
COMMENT Town,C.D., Van Aken,S., Uteback,T., Koo,H. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
source
1. 804
Location/Qualifiers

/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOHR746"
/note="Vector: PHOS1; Site 1: BstXI, 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 241 a 182 c 155 g 226 t

QY 995 AAGTAAACCGGTTGATCCCGATTATATC 1026
DB 545 AAGTAAACCGGTTGATCCCGATTATATC 514

Search completed: January 16, 2004, 19:16:55
Job time : 4299 secs

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rn1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 01:32:45 : Search time 104 Seconds
(without alignments)
6760.801 Million cell updates/sec

Title: US-10-022-025a-2
Perfect score: 1593
Sequence: 1 atgctcgaactcgaag.....tgcgcctagcgcagttaa 1593
Scoring table: IDENTITY_NTC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgm2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgm2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgm2_6/ptodata/1/ina/5A_COMB.seq:*
4: /cgm2_6/ptodata/1/ina/5B_COMB.seq:*
5: /cgm2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgm2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	651.6	40.9	1611	3	US-08-948-564-7 Sequence 7, Appl1
2	355	22.3	1799	3	US-08-560-398-3 Sequence 3, Appl1
3	144.2	9.1	1704	3	US-08-948-564-17 Sequence 17, Appl1
4	129.6	8.1	1708	3	US-08-991-677-1 Sequence 1, Appl1
5	98.2	6.2	275	4	US-09-313-294A-2591 Sequence 2591, Ap
6	93.8	5.3	1788	3	US-08-948-564-9 Sequence 9, Appl1
7	93.6	5.9	1893	1	US-08-532-065B-1 Sequence 1, Appl1
8	88.6	5.6	1761	3	US-09-033-055A-3 Sequence 3, Appl1
9	86.8	5.4	2174	3	US-08-606-505B-63 Sequence 63, Appl1
10	86.8	5.4	2174	3	US-09-616-990-63 Sequence 63, Appl1
11	86.2	5.4	1927	3	US-08-606-505B-64 Sequence 64, Appl1
12	86.2	5.4	1927	3	US-09-616-990-64 Sequence 64, Appl1
13	79.6	5.0	1657	3	US-08-948-564-11 Sequence 11, Appl1
14	78.6	4.9	1835	4	US-09-564-808-5 Sequence 5, Appl1
15	78.6	4.9	1838	4	US-09-091-432-1 Sequence 1, Appl1
16	78.6	4.9	1838	4	US-09-387-663-1 Sequence 1, Appl1
17	77.2	4.8	5156	2	US-09-091-432-3 Sequence 3, Appl1
18	77.2	4.8	5156	2	US-09-387-663-3 Sequence 3, Appl1
19	75.6	4.7	384	4	US-09-615-192A-232 Sequence 232, App
20	75.4	4.7	1722	3	US-09-033-055A-1 Sequence 1, Appl1
21	75.4	4.7	1735	1	US-07-912-900-29 Sequence 29, Appl1
22	75.2	4.7	1755	1	US-08-285-309-29 Sequence 29, Appl1
23	75.2	4.7	1755	1	US-08-502-046-29 Sequence 29, Appl1
24	75.2	4.7	1812	1	US-08-313-075A-37 Sequence 37, Appl1
25	75.2	4.7	1824	3	US-08-606-505B-1 Sequence 1, Appl1
26	75.2	4.7	1824	3	US-09-616-990-1 Sequence 1, Appl1
27	74.4	4.7	1620	4	US-09-627-216A-11 Sequence 11, Appl1

28	74.4	4.7	1620	4	US-09-765-873A-11 Sequence 11, Appl1
29	73.8	4.6	1884	4	US-09-564-808-3 Sequence 3, Appl1
30	72.8	4.6	543	4	US-09-615-192A-230 Sequence 230, App
31	72	4.5	1737	4	US-09-126-420A-1 Sequence 1, Appl1
32	70.6	4.4	1880	4	US-09-564-808-1 Sequence 1, Appl1
33	70.6	4.4	1634	4	US-09-126-420A-2 Sequence 2, Appl1
34	67	4.2	1665	1	US-08-313-075A-29 Sequence 29, Appl1
35	66.6	4.2	1824	3	US-08-948-564-13 Sequence 13, Appl1
36	65.2	4.1	1866	4	US-09-615-192A-103 Sequence 103, App
37	65.2	4.1	2013	4	US-09-615-192A-404 Sequence 404, App
38	64.4	4.0	1506	3	US-09-158-767-7 Sequence 7, Appl1
39	64.4	4.0	1506	3	US-09-158-767-8 Sequence 8, Appl1
40	64.4	4.0	1506	3	US-09-158-767-9 Sequence 9, Appl1
41	64.4	4.0	2261	3	US-09-351-229-3 Sequence 3, Appl1
42	64.2	4.0	1806	4	US-09-561-229-3 Sequence 3, Appl1
43	64.2	4.0	1812	1	US-07-912-900-28 Sequence 28, Appl1
44	64.2	4.0	1812	1	US-08-285-309-28 Sequence 28, Appl1
45	64.2	4.0	1812	2	US-08-502-046-28 Sequence 28, Appl1

ALIGNMENTS

RESULT 1
US-08-948-564-7
Sequence 7, Application US/06948564
Patent No. 6121512
GENERAL INFORMATION:
APPLICANT: Simlinszky, Balazs
APPLICANT: Dewey, Ralph E.
APPLICANT: Cordin, Frederick T.
TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512th Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948, 564
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCE/DOCKET NUMBER: 5051-409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-854-1400
TELEFAX: 919-854-1401
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1611 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 20..1588
US-08-948-564-7
Query Match 40.9%; Score 651.6; DB 3; Length 1611;
Best Local Similarity 66.5%; Pred. No. 2.9e-218; Indels 33; Gaps 5;
Matches 1025; Conservative 0; Mismatches 484;

QY 61 AGCCAAACCACTTGGCTTCTCCCTCCGCTGCAATCATCTGGCTGCGCATCT 120
 Db 74 ACCCAAGAAACCTTGGATGGGCTCTTTTGTATGATGGGCTCATCTGTTAACATGACT 133
 QY 121 CTCTTCTTATGACCTTATCCCGGTGACCTGCTGGGGAAATCCTCTTCCGCGGTTA 180
 Db 134 TTTATTAATCTGTACACCCGCTGGTGTCTGCGGCAAGTACTAACCC-----TA 186
 QY 181 ATATCCGGTATACAAACCGGAAACGTTATTCGCGTCCAAAGGCTTCCCTTGGTT 240
 Db 187 CTCTCCCCCTTTCAA-----TCAATCCCGGTCCCAAGGCTTCCCTTATT 235
 QY 241 GGAAGCATGCTACTATGTCAGACCTTAAGTCAACGAAATGCTGATGACGCTGAG 300
 Db 236 GGAACCATGGGCTTCATGACTTCC---CTGGCCCATCACTGATGCAAGCGCGCGCC 292
 QY 301 AAATCGAGCCCAAGAGCTCATGGCTTTTCACTTGAAGAGACTCGCGTATGCTACG 360
 Db 293 AATGCAAGAGCCAGAGCCCTCATGGCTTTAGTCTCGGCGCAACACGCTGATGCTACG 352
 QY 361 TGCATTCGCGAGTGAAGCAAGATTTCTGAATAGCCCGGTTTTTGTGATGACCGGTT 420
 Db 353 TGCCACCCCGAGTGGCCAGAGAGATTTCTCAACAGCTTCGCTTCCGCGATGCTCCGTC 412
 QY 421 AAAGAAATCGGCTTACTCACTGATGTTTAAAGAGCAATGGTTTTGACACACAGGTT 480
 Db 413 AAAGAAATCGGCTTACTCACTGATGTTTAAAGAGCAATGGTTTTGACACAGGTT 472
 QY 481 TAATGCGCAACGCTTCGCGGATGCTTGAACCATCTCTTTAGTACAAACAAATGCA 540
 Db 473 TAATGCGCAACGCTTCGCGGATGCTTGAACCATCTCTTTAGTACAAACAAATGCA 532
 QY 541 AGAGCCGAGCCGCAAGAGTGAATCTCAAGCCGAGTGTGATGTTCTTGAACAAAG 600
 Db 533 GCTTGTGAGCTCCAAAGCTCTCAAAATGCGCGCCCAAAAGTTCACATCT---AAATAC 589
 QY 601 AGTATGAACGAACCTGTTTTTGTGATGTTTAAAGCGCGCTGCTTAACACATG 660
 Db 599 AAGCCGACCGCAGCTTACGTGTGCGCAAGTGTGAAGAGCTTCCTCAGTAACTG 649
 QY 661 ATGTGCTGTATTTGCGCAAGATGATGCTGAAAAACCA-----TGTGAG 711
 Db 650 ATGTGCTGTATTTGCGCAAGATGATGCTGAAAAACCA-----TGTGAG 709
 QY 712 TTACGTGAATGATGCGAAGATGATGTTTGTGCGAAGCTGTAATGATGATGATC 771
 Db 710 CTTGGAATATGATGCGAAGATGATGTTTGTGCGAAGCTGTAATGATGATGATC 769
 QY 772 CTTCTTGGCTATCGAGTTTATCTCTCAAAAGCTCGGTCTGATGTTTCCACACTGTA 831
 Db 770 CTTCTTGGCTATCGAGTTTATCTCTCAAAAGCTCGGTCTGATGTTTCCACACTGTA 829
 QY 832 CCAAAGTGAACCGGTTGTATCCGGAATATATCGAAGCCGTAATCAACCGGTTAT 891
 Db 830 CCAAAGTGAACCGGTTGTATCCGGAATATATCGAAGCCGTAATCAACCGGTTAT 889
 QY 892 TTGCTGTGATTTGCTGAGCTTTTGTCTCTCTCCCTCACTGTTGATGATGATGATC 951
 Db 890 ACCAATCGGATTTTGTGATGCTGCTCTCTCTCCCGGAACCTGATCAATTAACAGAC 949
 QY 952 CCGGACATATCGCGCTCTTTGGGAGATGATTAATCAAGAGACAGACAGTTCGCTC 1011
 Db 950 TCCGACATATCGCTTACTTTGGGAATGATTAATCAAGAGACAGACAGTTCGCTC 1009
 QY 1012 TTAATGAGTGAATCTCTGCTGAGATGCTCTTCAATCGAATATCAACGTTACAA 1071
 Db 1010 TTAATGAGTGAATCTCTGCTGAGATGCTCTTCAATCGAATATCAACGTTACAA 1069
 QY 1072 AACGAGCTGATCAATGATGCGGAAATCAAGAGCCCTGATGATGATGATGATGATC 1131
 Db 1070 GAGAGCTGATGATGATGATGCGGAAATCAAGAGCCCTGATGATGATGATGATGATG 1129
 QY 1132 CTTCAATATCAAGGCTGTGTGAAAGATTAAGGCTTCAATCTTCCAGGCCACTT 1191

Db 1130 ATGAGTACTTACACGCGGTGTGAAGAGGTCTGCGGCTGCAACCCCGGCGCACTT 1189
 QY 1192 CTATCATGGGCGCGTTTGGCCATAACAGACAGATGTTGATGCTGCTGTTTCCGCA 1251
 Db 1190 CTATCATGGGCGCGTTTGGCCATAACAGACAGATGTTGATGCTGCTGTTTCCGCA 1249
 QY 1252 GGAACCAAGCATATGATGATGATGCTGCGGCGGTATGATGATGATGATGATGAT 1311
 Db 1250 GGAACCAAGCATATGATGATGATGCTGCGGCGGTATGATGATGATGATGATGATG 1309
 QY 1312 CTTTGAAGTATTAACCTGAGAGTCTGTCGCAAAAGAGTGAAGTGAAGTTCGCTT 1371
 Db 1310 CCACTGAAATTAATGCTCGAGAGGTTTGTCACTGCGGCGTGAAGATGCCGAATTTT 1369
 QY 1372 CTTGGGTGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1431
 Db 1370 CTTGGGTGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1429
 QY 1432 AATCTGCTTACTACGCTTATGTTTGAAGAGGATGATGATGATGATGATGATGATG 1491
 Db 1430 ACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1489
 QY 1492 GAGCCGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1551
 Db 1490 GATCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1549
 QY 1552 GCTAATCTCTTCTGCTAATTTGGCGGCTAGCGCGATTA 1593
 Db 1550 GCTAATCTCTTCTGCTAATTTGGCGGCTAGCGCGATTA 1591

RESULT 2
 US-08-560-398-3
 ; Sequence 3, Application US/08560398
 ; Patent No. 5907082
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Neill, Sharmen
 ; APPLICANT: Nadeau, Jeanette
 ; TITLE OF INVENTION: Oocyte-Specific Gene Expression
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/560,398
 ; FILING DATE: 17-NOV-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Baetian, Kevin L.
 ; REGISTRATION NUMBER: 34,774
 ; REFERENCE/DOCKET NUMBER: 023070-063300US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1799 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: unsure

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rnt

Page 3

LOCATION: 19..23
FEATURE:
NAME/KEY: CDS
LOCATION: 291..1571
OTHER INFORMATION: /note= "clone 040 ovule-specific gene
OTHER INFORMATION: encoding a cytochrome P450 monooxygenase
OTHER INFORMATION: from pollen tubes of Phalaenopsis"
US-08-560-398-3

Query Match 22.3%; Score 355; DB 2; Length 1799;
Best Local Similarity 58.0%; Pred. No. 6,4e-114;
Matches 744; Conservative 0; Mismatches 500; Indels 39; Gaps 5;

QY 304 TTGAGAGCCAAAGAGCTCATGCTTTCACTTGAAGAGACTGCGGTGATGCTACAGTGC 363
DB 273 TTGAGAGCCCTCCGCTCATGCTTTCTGCGTCCGCTCATGCTTCACTGCTTCAAGC 332
QY 364 AATCCGAGCTAGCGAAAGATTCGAATAGCCGCTTTTGTGATGACCGGTAA 423
DB 333 CACCCGAAACCGCAAAAGATTTCTTCAAGCCGCTTGCCTGATGCGCCATTAA 392
QY 424 GAATCGCTTACTCATGATGTTTAAAGAGCAATGCTTTGACACAGCGGTGTTAC 483
DB 393 GAATCGATACGAACCTTCTTTAAACGCGTATGGGTTTGGCCCATTTGGGAGTAC 452
QY 484 TGGCGAAGCTTCCGCTATGCTTGAACCATCTCTTAGTACAAAACAAATCAGAGA 543
DB 453 TGAAGAAACCTGAAAGATTTGTCACATCTTTTCACTCCGCGGAGTTTCATCG 512
QY 544 GCCGAGCGCAACGAGTATCTCAAGCCAGATGCTTGAATTTCTTGAAGAAACAGA-- 601
DB 513 TTGAGAGAGCAAGAGTGTGATTTGGGAAAGAAATGCTGGGATTTGAAGAAATGATG 572
QY 602 -GTATGACGAACCTGCTTTTGTCTGATGCTTAAACGCGCTGCTTAACATAG 660
DB 573 GAGAGAAATGAGATTTGTAAGAGTGAAGAGAAATTTCACTACGCGGCTTTGAATACATC 632
QY 661 ATGAGCTGTATTCGGAAGAAGTATGAGCTTGAAGAAACCATGCTGATTAAGTAA 720
DB 633 ATGATGCTGTTTTGGAGAAAGTTGATTTTGAAGAAAGATGAGGCTTGAAGCTTGA 692
QY 721 ATGCTC--GAAAGAGTATGATTTGCTCGAAAGCTTGAATGATGATCACTCTCT 777
DB 693 TTGATCTTAAAGAAAGATTTGATTTGCTTGAATCTTCACTGCGGCTGATCTTTGCT 752
QY 778 TGGCTATCGAGTTTATCTTCAAGACTCGGCTATGATTTCACTCGTACCAAG 837
DB 753 CTTTGGAGATGTTAGTTTGAAGGTTGTGAAGAGAAATGACGAACCTTGGCTTAAG 812
QY 838 GTAACCGCTTTGATCCGAGTTATATCCGAACACCGTATCAAAAC----- 885
DB 813 GTCAATGATTTTGAAGAAAGATCATAGACAGATTAAGAGAGCAACGCGTAAAG 872
QY 886 -----GCTATTTGCTCGTATTTGCTGACGTTTGTCTTCCCTCAATGAT 939
DB 873 ATTGATAGGATGAAGTGAAGATTTGATGCTTTGTTGGAGGAGAAAGAT 932
QY 940 AATATTCGACCCGCAATATGCGCTTTTGGAGATGATATTCAGAGAAACGAG 999
DB 933 AGATCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 992
QY 1000 AAGATTTGCTTTAATGAGTATCTGCTGATGATGATGATGATGATGATGAT 1059
DB 993 ACTGTTGCCATTCATTTGATGATGATGATGATGATGATGATGATGATGAT 1052
QY 1060 TCAACGATCAAAACGAGCTGATCAAGTATCGGAAATCAAGAGCCATGATGAT 1119
DB 1053 TCGAAGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1112
QY 1120 GACTTGGCTCACTTCATATCTAACGCTGCTGATGATGATGATGATGATGAT 1179
DB 1113 GATATCCAGCACTTCTTATCTCAATCTATAGTAAAGAAACCTTCAGATGATGAT 1172

QY 1180 CCAGGCCCACTTATATATGAGGCCGTTTGGCCATAAGACAGATGCTGATGCT 1239
DB 1173 CCTGGGCTCTATATGATGATGAGGCTGCGCTAGCTATCATGATGCTTCTGTTGATGCTAC 1232
QY 1240 CTGTTCCGAGAGGAGCAACAGCAATGATGATGATGATGATGATGATGATGAT 1299
DB 1233 ATGATTTCTGCTGAGAGATGATGATGATGATGATGATGATGATGATGAT 1292
QY 1300 GTGAGGTTATCTTTGAGTTTAACTGAGAGGTTGCTGCAAAAGAGTGAAGCT 1359
DB 1293 AACTGGCTGAGCTTCAAAATTCATCTGATGATGATGATGATGATGATGATG 1344
QY 1360 GAGTTTCCGTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1419
DB 1345 ----TCAATATCTTGGTTGATTTAAGTTGAGCTTGGCTCCGTTAAAGAT 1400
QY 1420 TGCCCGGAGAAATCTTGTGTTTACTACCGTTATGTTTGAACGCGGATGATGAT 1479
DB 1401 TGCCCTGCAAAAGATGAGCTTGTGATGATGATGATGATGATGATGATGAT 1460
QY 1480 GAGTTTGAATGAGAGCGCTCGAGTGAAGCGCTTACTATCTGAGAAACGAGCT 1539
DB 1461 AGCTTCAATTTGCTTCTTC--GAGAAATGATGATGATGATGATGATGATG 1517
QY 1540 TCTTGCAGATGCTTAATCTCT 1562
DB 1518 TCTCTGAGATGAAGATCTTT 1540

RESULT 3
US-08-948-564-17
Sequence 17, Application US/08948564
Patent No. 6121512
GENERAL INFORMATION:
APPLICANT: Siminszky, Balazs
APPLICANT: Dewey, Ralph E.
APPLICANT: Corbin, Frederick T.
TITLE OF INVENTION: No. 6121512a1 Cytochrome P-450 Constructs and
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512th Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,564
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCE/DOCKET NUMBER: 5051-409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-854-1400
FAX: 919-854-1401
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 38..1564

US-08-948-564-17

Query Match 9.1%; Score 144.2; DB 3; Length 1704;
 Best Local Similarity 53.5%; Pred. No. 8.3e-40;
 Matches 355; Conservative 0; Mismatches 293; Indels 15; Gaps 2;

QY 866 CCGAAGACCGTAAATCAACCGGTGATTTGCTCGTATTTTCGTGACGTTTGTCTCC 925
 DB 813 CTGAGACACGCAAAATCTGTGTGTCAGAACAACTTTGTGATGCTCTCCACAT 872
 QY 926 TCCATGCTCAGATAATTAATCCGACCGGACATATCGCGTCTTTGGGAGATGAT 985
 DB 873 TCGAAGACAATATGACCTTAGTAGAGACACATCATTTGCTCTTTGGGATATGATCA 932
 QY 986 TCAAGAGACACACACACTTCCGCTCTTAATCGAGTGATCTCTGTAAGATGCTTC 1045
 DB 933 CAGCAGGATGACACACACTGCAATTCAGTTGAGTGGCCATGGCTGAGTTGATAGAA 992
 QY 1046 ATCCAGATATGCAATCAACGGTACAAACGAGTCGATCAAGTAGTCGGGAAATCAAG 1105
 DB 993 ACCCAAGGTCAACAAAGTCTCAAGAGCTAGACAGGTATATGGCTTGAAGGG 1052
 QY 1106 CCTAGATGAATCTGACTTGGCTTCACTTCATATCTTAACGGCTGTGTGAAGATAT 1165
 DB 1053 TGATGACTGAGCAGACTTCTCAATCTCCTTACTTCAATGTGTACCAAGAACAA 1112
 QY 1166 TGAAGCTTCACTCCACAGGCCACTTCTATCATAGGGCCGTTGGCCATAACAGACGA 1225
 DB 1113 TGAGCTTCCACCAACCCCACTAATGATCTCCACACCGT---GCCATGCAATGTCA 1169
 QY 1226 TCGTATGATGTCCTTCTTCCGCGACGAGACACAGCATGTGAACATGTGGCCGTAT 1285
 DB 1170 AAGTTGAGGCTATGACATTCCTCAAGAGGTCCAAATGTGATGATATGTGGCGGTGG 1229
 QY 1286 CGCATATCCACACGTGTGGGTGATCTCTTGGAGTTAACTGAGAGGTTCTGTGCA 1345
 DB 1230 CCGGACCGCGCGCTGTGAGAGATCCATGAGATCCGACCGGAGGTTCTTTAGG 1289
 QY 1346 AAGAGGTGAGGTGAGTTTCGATTTCTTGGGTGCGATTTGAGACTTTCGAGT 1405
 DB 1290 AGGATGTAGACATGAAG-----GGCATGACTTAAAGCTTACATTCAGGT 1337
 QY 1406 CGGTCGTGCGATTTGCCCCCGGAGAAATCTTGGTTTACTACCGTATATTTTGACGG 1465
 DB 1338 CGGTCGACGAGTATGCCCCGGGTGCCCACTTGTATCACTTGGACACATCCATGTGG 1397
 QY 1466 CGATGATGTACATGATTTGATGGGACCGTCGATGTAAAGCGCTTGAATTTATCTG 1525
 DB 1398 GGCACCTCTTGACACATTTCTGTGTGACCCCACTGAGAGAAATGAAGCTGAGAAATG 1457
 QY 1526 AGA 1528
 DB 1458 ACA 1460

RESULT 4

US-08-991-677-1
 ; Sequence 1, Application US/08991677A
 ; Patent No. 6252135
 ; GENERAL INFORMATION:
 ; APPLICANT: Chiang, Vincent L
 ; APPLICANT: Carraway, Daniel T
 ; APPLICANT: Smeltzer, Richard H
 ; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
 ; FILE REFERENCE: 50617
 ; CURRENT APPLICATION NUMBER: US/08/991,677A
 ; EARLIER FILING DATE: 1997-12-16
 ; EARLIER APPLICATION NUMBER: US 60/033,381
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1708

TYPE: DNA
 ORGANISM: Liquidambar styraciflua
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (48)..(1571)
 US-08-991-677-1

Query Match 8.1%; Score 129.6; DB 3; Length 1708;
 Best Local Similarity 52.8%; Pred. No. 1.1e-34;
 Matches 335; Conservative 0; Mismatches 284; Indels 15; Gaps 2;

QY 873 CCGTATCAAAACCGGTGATTTGCTCGTATTTGCTGACGTTTGTCTCTCCATAG 932
 DB 827 CCGTAAAAAGATGTGTGGGCCCAACACATTTCTGTGATCATTTGCTCACCTTACAGA 886
 QY 933 TTCAATTAATTAATTCGACCGGACATATATCGCTTTTGGGAGATGATATTCAGAG 992
 DB 887 GAATATGACTTAACTGAGAGACATATATTTGGGCTCTTTGGGATATGATCATGACAG 946
 QY 993 AACAGACAGTGTGGCTTTATCGAGTATCTCTGATGATGTCCTTCAATCCAGA 1052
 DB 947 CATGACACACCGCAATCTGTGAAATGGCCATGCGCGATTAATTAAGAACCAAG 1006
 QY 1053 TATGCAATCAACGTAACAAACGAGCTGATCAATGATCGGAAATCAAGCCCTTGA 1112
 DB 1007 GGTCACAAACAAAGCTCAAGAGAGCTAGACATGATGATTTGGTCCGAACGTGCTTAC 1066
 QY 1113 TGAATCTGACTTGGCTTCACTTCATCTTAACGGCTGTGTGAAGATTTAGGCT 1172
 DB 1067 CGAATGACTTCTCAAGCTCTCTTATCTAATGTGTAGCAAGAGGACATAAGCT 1126
 QY 1173 TATCTCTCAGGCCCACTTATCATAGGCGCTTGGCCATTAACAGACAGATGTTGA 1232
 DB 1127 GACCTCTCA---ACACCACTATGCTCTCTATGCGCGCATATGCAACGTCAAAATGG 1183
 QY 1233 TGGTGTCTTGTTCGCGGACGGGACACAGCATGTGGAACATGTGGCGTATCGCATGA 1292
 DB 1184 TGGCTACACATCTCTTAAGGATCAAAATTTATGATTAATGTCTGGCGGTGCTGTGA 1243
 QY 1293 TCCACAGTGTGGGTGATCTCTTGAAGTTAAACCTGAGAGTGTGTGGCAAGAGG 1352
 DB 1244 TCCAGAGATGTGGCGTGAACCACTAGATTTGACCGGAAACGTTCTCTGAAGAGATGT 1303
 QY 1353 TGAAGTGAATTTGCGTCTTGTGCGATTTGAGACTTGACCTTTCGAGTGGGTGCG 1412
 DB 1304 CGACATGAAA-----GATCAGATTAATAGCTACTGCGGTTTGGTCAAGGAG 1351
 QY 1413 TCGATTTGCCCCCGGAGAAATCTTGTTTTACTACCGTTATGTTTGGACGGCATGAT 1472
 DB 1352 GCGTGTGTGGCGCGGTGACCAACTTGACATTTGTGTGACATCATGATGAGTCACT 1411
 QY 1473 GTTACATGATTTGAATGGGACCGTCCCATGT 1506
 DB 1412 ATTGACCAATTTCTATTTGAGCCCTCTAAAGGT 1445

RESULT 5

US-09-313-294A-2591
 ; Sequence 2591, Application US/09313294A
 ; Patent No. 6476212
 ; GENERAL INFORMATION:
 ; APPLICANT: Lalundi, Raghunath V.
 ; APPLICANT: Ito, Laura Y.
 ; APPLICANT: Sherman, Bradley K.
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
 ; FILE REFERENCE: PI-0017 US
 ; CURRENT APPLICATION NUMBER: US/09/313,294A
 ; EARLIER FILING DATE: 1999-05-14
 ; NUMBER OF SEQ ID NOS: 7600
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 2591
 ; LENGTH: 275
 ; TYPE: DNA

```

: ORGANISM: Zea mays
:
: FEATURE:
: NAME/KEY: misc feature
: OTHER INFORMATION: Incyte ID No. 6476212 700552946H1
US-09-313-294A-2591

```

Query Match	6.2%;	Score 98.2;	DB 4;	Length 275;
Best Local Similarity	64.7%;	Pred. No. 3.4e-24;		
Matches 178; Conservative	0;	Mismatches 93;	Indels 4;	Gaps 2

OY	1199	GGGGCCCGTTTGGCCCTTAACAGACAGATCGTTGAGTGGTCTGTCTGTCTTCCGACAGAACCA	1259
Db	5	GGGGCGCGCTCGCCACCTCAGACGACAGTAAGGGGGTACTTGTGTGCGCGGACCA	64
OY	1259	CAGCAATGATGAAACATGTGTGGCGCTATGCAATGATCCAACAGTGTGGTGTGATCTTTGG	1318
Db	65	CCGCATATGTGAACATGTGTGGCCATTAACCAATGACCCAGCTGTGGCTGTGACCAATGG	124
OY	1319	AGTTTAAACCTGAGAGTTCGTGGGCAAAAGAGTGAAGATTTCGTCTTGGGT	1378
Db	125	AGTTACAGCCCGAAGAGTTTCAATGGCCCTGCGCCGCA---GGACCTCCCGAATAAGGGT	181
OY	1379	CGAATTTGAGACTTCACCTTTGGGGTGGGGTCGTGGATTTTGCCCGGGAAGATCTTG	1438
Db	182	CGGACTTCGGCTCCGCGCTTTTGGGGTCCGGAGGCGGA-GTGCCCGCGCAAGTCACTCG	240
OY	1439	GTTTACTACCGTTATGTTTGGACGGCGATGATG	1473
Db	241	CGGTGGCTACCGTGGATTTCTGGTCCGCACTTG	275

RESULT 6
US-08-948-564-9
Sequence 9, Application US/08948564
Patent No. 6,215,152
GENERAL INFORMATION:
APPLICANT: Siminszky, Balazs
APPLICANT: Dewey, Ralph E.
APPLICANT: Corbin, Frederick T.
TITLE OF INVENTION: No. 6,125,151
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:

NAME/KEY:	CDS
LOCATION:	6..1601
US-08-948-564-9	

Query Match	5.94;	Score 93.8;	DB 3;	Length 1788;
Best Local Similarity	48.0%;	Pred. No. 4.6e-22;		
Matches 303; Conservative	0;	Mismatches 322;	Indels 6;	Gaps 1

QY	933	TTCCGATTAATAATCCCAACCCGGACATTAATCCGCGCTCTTGGAGATGTATTCAGAG	992
Db	932	TTCTGGTTATGATTCAGATACCATCAATCAAGGCTACTGTCTGATATCTGATTTTAGCAG	991
QY	993	AACAGACACAGTTGCGGTCTTAAACGAGTGAATCCCTCGTAGATGTCCTTATCCAGA	1052
Db	992	AAGCGACACCAACATGATTTCACTAAATGAGTGTCTCTCTGTACTTAAACATCAAT	1051
QY	1053	TATGCAATCAACGGTACAAAACGAGCTGATCAAGTAGTGGGAAATCAAGAGCCCTAGA	1112
Db	1052	GGAACCTAAAAAAGTCCAAAGTAGATTTGGACACTTATATTTGGGAAGGACGAAAGTGGa	1111
QY	1113	TGAATCTGACTGGCTTCACTTCATATCTAACGGCTGTGATGAAGAGATTAAGGCT	1172
Db	1112	AGAAATCTGACATAACCAAGTTGGTGTACCTCCAGGCATTTGTGAAGGAAACAATGGGCT	1171
QY	1173	TCATCTCCAGAGCCCACTTCTATCATGAGGCCCGTTGGCCATTAACAACACGATCTGTGA	1232
Db	1172	GTATCCACCAAGTCTCTTATCAACCTTCGTGACGACATGGMAACATGCACTTCTCAGG	1231
QY	1233	TGGTGTGTGTATGTCGGGACAGGACACAGCATGTGTAACATGTGGGCGGTATCCAGTA	1292
Db	1232	TGGCTATCACTTCTGTGTSBGAACAGTTTAAATGTGAATCTTGAAGATCCACCGGGA	1291
QY	1293	TCACACGTGTGGGTGAATCTCTTGGAGATTAAACCTGAGAGTTTCGTGGGAAAGAGG	1352
Db	1292	TGGTGTGTGTATGAGATGATCTCATGATTTAAACCTGGAAGGTTCTTGAACAACCA	1351
QY	1353	TGAGGTGAGATTTCCGATCTTCTTGGAGTCGAGATTGAGACTTGACCTTTCCGGTCCGCTG	1412
Db	1352	AGATGTGATGGAAGGGTCAAGAACTATAG-----CTCTCTCTTTTGGTTCTTGGAAG	1409
QY	1413	TCCGATTTGCCCGGGAGAGATCTTGGTTTACTACGTTATGTTTGGACGGCGATAT	1472
Db	1406	GAGACGATGCCCTGAGCCTCGCTGGCTCTGCGTGTGTTGACACTTGAACAGTGGCTAGCT	1465
QY	1473	GTTACATGAGTTGAATGAGGAGCCGTCGAGTGTAAACGGCGTTGACTTATCTGAGAACT	1532
Db	1466	GTTACATCTTCAATGTTGTCTTCTCTTCAAACTCAAGTTGTGACATGAACGAGAGCAT	1525
QY	1533	GAGGCTTTCTGCGAGATGGGTAATCCTCTT	1563
Db	1526	TGACATCAAAATTTAAAGCAACCCCGCTT	1556

RESULT 7
 US-08-532-065B-1
 Sequence 1, Application US/08532065B
 Patent No. 5753507
 GENERAL INFORMATION:
 APPLICANT: Ohta, Daiaku
 APPLICANT: Mizutani, Masaharu
 TITLE OF INVENTION: plant Gellanol/Nerol 10-Hydroxylase and
 TITLE OF INVENTION: DNA Coding Therefor
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: No. 5753507artis Corporation
 STREET: 59 Route 10
 CITY: East Hanover
 STATE: NJ
 COUNTRY: USA
 ZIP: 07936
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/532,065B
 FILING DATE: 22-SEP-1995
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Meigs, J. Timothy
 REGISTRATION NUMBER: 38,241
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8587
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1893 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 ORIGINAL SOURCE:
 ORGANISM: Arabidopsis thaliana
 IMMEDIATE SOURCE:
 CLONE: P450-4
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 6..1490
 US-08-532-065B-1

Query Match 5.9%; Score 93.6; DB 1; Length 1893;
 Best Local Similarity 50.7%; Pred. No. 5,6e-22;
 Matches 287; Conservative 0; Mismatches 264; Indels 15; Gaps 2;

941 AATTATCCGACCCGGACATATATCGCGTCTTTGGGAGATGATATTCAGAGAACAGACA 1000
 889 AACTCTCATTAGCATATGAAACACTTCTTGATGATGTTTACAGACAGACAGATA 948
 1001 CAGTGGCGGTCTTATTCAGAGATCCGCTAGAGATGATGATGATGATGATGAT 1060
 949 CAGCTCTTATGTCCTGAGTGCCCAATGACAGATTTCTTAAACCTTAAACGATG 1008
 1061 CAACGGTACAAACGAGTGAATCAAGTATGCGGAAATCAAGACCTTATGATGATG 1120
 1009 CAAAAGCTCAGCCGAGATCGATGATGATGATGATGATGATGATGATGATGATG 1068
 1121 ACTTGGCTTCACTTCCATATCTTAAGCGGTGATGAAAGATATGAGGCTTATCCTC 1180
 1069 AATATCTCAAACTGCGTATTTTCAAGACGATGATGAGAACTTTCCGCTTACATGCG 1128
 1181 CAGGCCACTTTCATCATGAGGCGCTTTGCGCATTAACAGACAGATCGTTGATGCTG 1240
 1129 CTTGTCGCTTCT---TATCCGCGAAAGCGAATCCGATCGGAGATTTCTTGGTTTCA 1185
 1241 TTGTTCCGCGAGGACCAAGCAATGATGATGATGATGATGATGATGATGATGATG 1300
 1186 TGTGCTTAAAGTATCTCAGGTTCTAGTGAAGCTTGGCCATAGACAGACCCGAGCG 1245
 1301 TGTGCGTTGATCTCTTGGATTAAACCTGAGGTTCTGTCGCAAAAGAGGTGAGGTG 1360
 1246 TGTGGATATATCGTCCAGTTTGAACAGAGAGGTTTTTGGGAAAGATATGACGTGA 1305
 1361 AGTTTTCGTTCTTGGGTGAGATTTGAGACTTGCACTTTCGGGTGCGGTGCGATGAT 1420
 1306 GAGTGAAGATTATGAG-----CTTACACATTCGCGCGGAGCGATGAAATTT 1353
 1421 GCCCGGAGAGATCTTGGTTTACTACGTTATGTTTGAACGCGGATGATGATGATGAT 1480
 1354 GCCCGGAGATCTTGGGTGATGAAAGAGTGTCTTATGCTTGTGCTTCTTATTT 1413
 1481 AGTTTGAATGAGGACCGTCCGATGCT 1506
 1414 CTTTGACTGGAAGCTTCCGAAAGGT 1439

RESULT 8
 US-09-033-055A-3
 Sequence 3, Application US/09033055A
 Patent No. 6069241
 GENERAL INFORMATION:
 APPLICANT: OKAMA, HIDEO
 APPLICANT: IMASHI, HIROMASA
 TITLE OF INVENTION: NOVEL CYTOCHROME P450 GENE
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PILLSBURY, MADISON & SUTRO
 STREET: 1100 NEW YORK AVENUE, N.W.
 CITY: WASHINGTON
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy Disk
 COMPUTER: IBM compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/033,055A
 FILING DATE: MARCH 2, 1998
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: PAUL E. WHITE, JR.
 REGISTRATION NUMBER: 32,011
 REFERENCE/DOCKET NUMBER: 9437/251563
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 861-3651
 TELEFAX: (202) 822-0944
 TELEX: 6714627 CUSH
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1761
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-09-033-055A-3

Query Match 5.6%; Score 88.6; DB 3; Length 1761;
 Best Local Similarity 48.8%; Pred. No. 3e-20;
 Matches 300; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

929 ATGATTCAGATTAATTAATCCGACCCGAGATTAATCGCGTCTTTGGGAGATGATTTCA 988
 923 ATGGAAACAAACATGATGACAGAGATTAAGCTTAATTCAGGATATGATGATGCTG 982
 989 GAGGACAGACACAGTGGCGGTCTTAATGAGTGAATCCGCTAGATGATGATGATGATG 1048
 983 CAGCCAGGATATATCTGCTGTCACCAACGATGCGCAATGCTTGGGATTAAGATTC 1042
 1049 CAGATATGATCAACGATCAAAAGAGTGAATCAAGTATGCGGAAATCAAGAGCCC 1108
 1043 CAATGCTCCAGAAAGATCAAGAAATTAATGATGATGATGATGATGATGATGATGAT 1102
 1109 TAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1168
 1103 TTAACGGAATCCGACTTGTTCATCTCAAGTACCTTGTGTGATGATGATGATGATGAT 1162
 1169 GGCCTTCACTCCAGAGCCCACTTCAATCATGAGGCGCTTTGGCCATTAACAGACGATCG 1228
 1163 GAATCACTCTGCTGCTCACTTTCT---AATTCACATGATCAATTCGATGATGATGAT 1219
 1229 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1288
 1220 TCAACGCTATTAATATCCAGAAAGACAGTGTCTTCATCAACACATGATGATGATGAT 1279
 1289 ATGATTCACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1345
 1280 GAAACCAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1339

QY 1346 AAGAGGTGAGGTGAGGTTTCGGTCTTGGGTCGAGTTTGAGACTTGACCTTGCGGT 1405
 DB 1340 ATGAATTAAGTAGAGTTGAATAATGATGAGGCGGAGATTTCAGATTTCACATTAGTG 1399
 QY 1406 CGGTCGTCGGAATTTGCCCCCGGAGAGATCTTGTTTATCTACCTTATGTTTGAACGG 1465
 DB 1400 CTGGAAGAAAGAGTGTCTGCTGTCACCAATTAGAGTGAATTTGGTACTTATGGCTTGG 1459
 QY 1466 CGATGATCTTACATGAGTTTGAATGGGACCGTCCATGTATACGCGCTTGAATTCTG 1525
 DB 1460 CTGAGTTGTTCCATGCTTGTGAGTGAGCCACCAAGATGATTAAGCGCTGAAGATATTG 1519
 QY 1526 AGAAGCTGAGGCTTT 1540
 DB 1520 AGACAATTGAGGTTT 1534

RESULT 9

US-08-606-505B-63

Sequence 63, Application US/0860505B

Patent No. 6114601

GENERAL INFORMATION:

APPLICANT: KIKUCHI, Yasuhiro

APPLICANT: KIYOKAWA, Shigeto

APPLICANT: SHIMADA, Yukihisa

APPLICANT: OHBAYASHI, Masaya

APPLICANT: SHIMADA, Ritsuko

APPLICANT: OKINAKA, Yasushi

TITLE OF INVENTION: NOVEL PLANT GENES

NUMBER OF SEQUENCES: 67

CORRESPONDENCE ADDRESS: 67

ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10112-3801

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.

COMPUTER: IBM PS/4

OPERATING SYSTEM: MS-DOS Ver3.30

SOFTWARE: PATENT AID Ver1.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/606,505B

FILING DATE: 23-MAR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP44963/92

FILING DATE: 02-MAR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Perry, Lawrence S.

REGISTRATION NUMBER: 31865

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-218-2100

TELEFAX: 212-218-2200

INFORMATION FOR SEQ ID NO: 63:

SEQUENCE CHARACTERISTICS:

LENGTH: 2174 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORGANISM: Eustoma russellianum

FEATURE:

NAME/KEY: CDS

LOCATION: 92 to 1621

IDENTIFICATION METHOD: by experiment

US-08-606-505B-63

Query Match 5.4%; Score 86.8; DB 3; Length 2174;
 Best Local Similarity 49.7%; Pred. No. 1.5e-19;
 Matches 307; Conservative 0; Mismatches 302; Indels 9; Gaps 3;

QY 956 ACATATCGCGCTTCTTGGGAGATGATATTCAAGAGAACAGACAGTTCGGCTTTAA 1015
 DB 972 ATATCAAGGCTCTTTTATTGAACATTTTACCGCTGTAGGATACATCTTCAAGGCTCA 1031
 QY 1016 TCGAGTGAATCTTCGTAGAGTGTCTTCATCCAGATATGACATCAACGTACAAAGC 1075
 DB 1032 TAGAGTGGGCGCTGGCCGAGTGTCTTAAGATCAATCATCTTAAGACGAGCCAGAAAG 1091
 QY 1076 AGCTGATTCAGATGTCGGGAAATCAAGAGCCCATAGTGAATGATCTGATGGCTTCACTTC 1135
 DB 1092 AATGACGGTGTATGCGCGAGACCGCGGCTTCTTGAAGCAGACATATCAAGTGC 1151
 QY 1136 CATATCTAAGCGCTGTGTGTAAGAAAGATATGAGGCTTCAATCTCCAGGCCCACTTCTAT 1195
 DB 1152 CATATCTCAAGCCATCTGCAAGAAAGCTTTCAGAAAGCATCTTC--CAGGCTTTAA 1208
 QY 1196 CATTGGCCCTTTGGCCATAACAGACAGATCGTGTATGTCGTCTTGTCCGGCAGGGA 1255
 DB 1209 ATCTCCACGATGCGCTCGCAGCATGTGAAGTAATGAACACTATACCAAGGCA 1268
 QY 1256 CCACAGCAATGATGACATGTGGCCGTATCGATGCCACAGTGTGGTGTATCTT 1315
 DB 1269 CTAGGCTCAGGCTTAACATATGGGCTATGTGAAGATCCATCTGTGTGGGAAATCCAA 1328
 QY 1316 TGGAGTTTAACTGACAGGTTCGTGGCAAAAGGTGAGTGAATTTTGGCTTCTG 1375
 DB 1329 ATGAATTTAACCCGATGAGTGTGTTTGAAGAAAGATGCCA--AGATGATCCACAG 1385
 QY 1376 GGTGGATTTGAGACTTGACACTTTCGGTGGGTCGTGGATTTGCCCGGAGAAATC 1435
 DB 1386 GAATGATTTTGAAGTATGCCATTTGGAGGTGAAGAAATTTGGCGTGAACAAGAT 1445
 QY 1436 TTGGTTTACTACCGTTATGTTTGGACGGGAGATGTTTACATGAGTTTGAATGGGA- 1494
 DB 1446 TGGGAATCTTCTAGTGAATATTTTGGAACTTTTGGTGCATCTTTTGTGGAAAT 1505
 QY 1495 --CGTCCGATGTATGAGCGCTTGAATCTGATGATGAGAGTTCCTTGGCAGATG 1552
 DB 1506 TGCCATCTCTGTGATTTGAATTTAATGATGATGATGCTTTTGGCTTCTGTGAGAGG 1565
 QY 1553 CTAACTCTCTTCTGCTA 1570
 DB 1566 CAGTCCCTCTGCTGCTA 1583

RESULT 10

US-09-616-990-63

Sequence 63, Application US/09616990

Patent No. 6232109

GENERAL INFORMATION:

APPLICANT: KIKUCHI, Yasuhiro

APPLICANT: KIYOKAWA, Shigeto

APPLICANT: SHIMADA, Yukihisa

APPLICANT: OHBAYASHI, Masaya

APPLICANT: SHIMADA, Ritsuko

APPLICANT: OKINAKA, Yasushi

TITLE OF INVENTION: NOVEL PLANT GENES

NUMBER OF SEQUENCES: 67

CORRESPONDENCE ADDRESS:

ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10112-3801

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.

COMPUTER: IBM PS/4

OPERATING SYSTEM: MS-DOS Ver3.30

SOFTWARE: PATENT AID Ver1.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/616,990

FILING DATE: 14-Jul-2000

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rml

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Perry, Lawrence S.
REGISTRATION NUMBER: 31865
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 2174 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Eustoma russellianum
FEATURE:
NAME/KEY: CDS
LOCATION: 92 to 1621
IDENTIFICATION METHOD: by experiment
SEQUENCE DESCRIPTION: SEQ ID NO: 63
US-09-616-990-63

Query Match 5.4%; Score 86.8; DB 3; Length 2174;
Best Local Similarity 49.7%; Pred. No. 1.5e-19;
Matches 307; Conservative 0; Mismatches 302; Indels 9; Gaps 3;
QY 956 ACATTAATGCCGCTTTTGGAGATGATATTCAGAGAAACAGACAGCTTGGCTTTAA 1015
DB 972 ATATCAAGGCTCTTTTATGAACATGTTTACCGCTGACGATACATCTTCAAGCTCA 1031
QY 1016 TCGAGTGAATCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1075
DB 1032 TAGAGTGGCGCTGGCCGAGATGCTTAAGAAATCCATCATCTTAAGACAGCCCAAG 1091
QY 1076 AGCTGATCAATGATCGGAATCAAGAGCCCTAGATGATGATGATGATGATGATGATGAT 1135
DB 1092 AATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1151
QY 1136 CATATCTCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1195
DB 1152 CATATCTCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1208
QY 1196 CATGGGCCGCTTTGGCCATAACAGACAGATGATGATGATGATGATGATGATGATGATGAT 1255
DB 1209 ATCTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1268
QY 1256 CCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1315
DB 1269 CTAAGCTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1328
QY 1316 TGAAGTTAAACCTGAGAGGTTCTGAGCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1375
DB 1329 ATGAATTTAACCTGATAGGTTTGAACGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1385
QY 1376 GGTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGAT 1435
DB 1386 GAAATGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGAT 1445
QY 1436 TTGATTTTACTACCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 1494
DB 1446 TGGGAATTTACTTGAATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 1505
QY 1495 --CCGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1552
DB 1506 TCCCATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1565
QY 1553 CTAATCTCTCTGCTGCTA 1570
DB 1566 CAGTCTCTTCTGCTGCTA 1583

RESULT 11
US-08-606-505B-64
Sequence 64, Application US/08606505B
Patent No. 6114601
GENERAL INFORMATION:
APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: SHIMADA, Yukihisa
APPLICANT: OHBAYASHI, Masaya
APPLICANT: SHIMADA, Ritsuko
APPLICANT: OKINAKA, Yasuhiro
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 KB storage.
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/606,505B
FILING DATE: 23-MAR-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Perry, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2200
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 1927 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Campanula medium
FEATURE:
NAME/KEY: CDS
LOCATION: 180 to 1748
IDENTIFICATION METHOD: by experiment
US-08-606-505B-64
Query Match 5.4%; Score 86.2; DB 3; Length 1927;
Best Local Similarity 49.9%; Pred. No. 2.3e-19;
Matches 273; Conservative 0; Mismatches 268; Indels 6; Gaps 2;
QY 959 TAATCCCGCTCTTGGAGATGATATTCAGAGAAACAGACAGATGATGATGATGATGATGATGATGAT 1018
DB 1102 TTAAGCACTCTCTTGGAGATGATATTCAGAGAAACAGACAGATGATGATGATGATGATGATGATGAT 1161
QY 1019 AGTGAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1078
DB 1162 AATGGGACATGAGCGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1221
QY 1079 TGAATGAT 1138
DB 1222 TGAACCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1281
QY 1139 ATCTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1198
DB 1282 ATTTCAAGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1338

```

QY 1199 GGGCCCGTTTGGCCATTAACAGACAGATCGTTGATGTCGTCCTTGTCCGCGAGGACCA 1258
DB 1339 TCCCAAGAAATCTCAACAGAGCATGTGAAGTGACCGAATTTCACTACCAAAAACACTA 1398
QY 1259 CAGCAATGGTGAACATGTGGGCGCTATCGCATGATCCACACGTCGTGGTGAATCTTGG 1318
DB 1399 GACTAATAGTGAACATGTGGGCAATGAGGAGGACCTTAAAGTGGGAAAATCCATTG 1458
QY 1319 AGTTTAAACCTGAAGGTCGTGCGCAAAAGAGTGAAGTGAATTTGCGTTTGGGT 1378
DB 1459 ATTTTACTCCGGAACGTTTCTTGAGTGAATAAAC--ACGGAAAATTTGATCCGCGAGTA 1515
QY 1379 CGAATTTGAGACTTTCACCTTTGGGGTGGGTCGTGGAATTTCCCGGGGAAAATCTTG 1438
DB 1516 ATCTTTTGAATTAATCCATTTGGGGCTGACGAAGATGTGCAAGGAGGCTGAGATGG 1575
QY 1439 GTTTTACTACCGTTATGTTTGGACGCGCATGATGTTACATGATTTGAATGGGACCGT 1498
DB 1576 GAGCGGCTCGTGGAGTACATATTAAGTATGATGTCGACCTGATTGATTTGAAATTC 1635
QY 1499 CCGATGG 1505
DB 1636 CTGATGG 1642

```

RESULT 12

US-09-616-990-64
Sequence 64, Application US/09616990
Patent No. 6232109

GENERAL INFORMATION:

APPLICANT: KIKUCHI, Yasuhiro
KIYOKAWA, Shigeto
SHIMADA, Yukihisa
OHBAVASHI, Masaya
SHIMADA, Ritsuko
OKINAKA, Yasushi
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSER: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/IV
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/616,990
FILING DATE: 14-Jul-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP44963/92

ATTORNEY/AGENT INFORMATION:

NAME: Perry, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200

INFORMATION FOR SEQ ID NO: 64:

SEQUENCE CHARACTERISTICS:
LENGTH: 1927 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Campanula medium
FEATURE:
NAME/KEY: CDS

```

; LOCATION: 180 to 1748
; IDENTIFICATION METHOD: by experiment
; SEQUENCE DESCRIPTION: SEQ ID No: 64
US-09-616-990-64

```

Query Match 5.4%; Score 86.2; DB 3; Length 1927;
Best Local Similarity 49.9%; Pred. No. 2, 3e-19;
Matches 273; Conservative 0; Mismatches 268; Indels 6; Gaps 2;

```

QY 959 TAATCCCGTCTTTGGAGATGATTAATCAAGAGACAGACACAGTTCGGTCTTAATCG 1018
DB 1102 TTAAGGCACTCTTTGGATTATTAATCAAGGCGGCGACGATATCATCAAGTGTATCG 1161
QY 1019 AGTGATCTCTGATGATGTCCTTATCATGATATGCAATCAAGGTACAAAGAGAGC 1078
DB 1162 AATGGGCACTAGCGAAATGTGAACCATGACAGATCTTAACCGGAGCCACGAAAGAA 1221
QY 1079 TGCATCAAGTATGTCGGGAATCAAGAGCCCTAGATGATCTGACTTGGCTTCACTTCAT 1138
DB 1222 TGAACCAAGTATTTGGGAGAAACAGAAAGACTAGAACATGTGACATACCAAACTTGCAT 1281
QY 1139 ATCTAAGGCTGTGTGTAAGAAAGATTAATGAGCTTCACTCCAGGCCCACTTATCAT 1198
DB 1282 ATTTCCAAAGCCATATGCAAGAAACATTCGAAACACCCCTTC--CAGCGCTTAACC 1338
QY 1199 GGGCCCGTTTGGCCATTAACAGACAGATCGTTGATGTCGTCCTTGTCCGCGAGGACCA 1258
DB 1339 TCCCAAGAAATCTCAACAGAGCATGTGAAGTGACCGAATTTCACTACCAAAAACACTA 1398
QY 1259 CAGCAATGGTGAACATGTGGGCGCTATCGCATGATCCACACGTCGTGGTGAATCTTGG 1318
DB 1399 GACTAATAGTGAACATGTGGGCAATGAGGAGGACCTTAAAGTGGGAAAATCCATTG 1458
QY 1319 AGTTTAAACCTGAAGGTCGTGCGCAAAAGAGTGAAGTGAATTTGCGTTTGGGT 1378
DB 1459 ATTTTACTCCGGAACGTTTCTTGAGTGAATAAAC--ACGGAAAATTTGATCCGCGAGTA 1515
QY 1379 CGAATTTGAGACTTTCACCTTTGGGGTGGGTCGTGGAATTTCCCGGGGAAAATCTTG 1438
DB 1516 ATCTTTTGAATTAATCCATTTGGGGCTGACGAAGATGTGCAAGGAGGCTGAGATGG 1575
QY 1439 GTTTTACTACCGTTATGTTTGGACGCGCATGATGTTACATGATTTGAATGGGACCGT 1498
DB 1576 GAGCGGCTCGTGGAGTACATATTAAGTATGATGTCGACCTGATTGATTTGAAATTC 1635
QY 1499 CCGATGG 1505
DB 1636 CTGATGG 1642

```

RESULT 13

US-08-948-564-11
Sequence 11, Application US/08948564
Patent No. 6121512

GENERAL INFORMATION:

APPLICANT: Siminszky, Balazs
APPLICANT: Corbin, Frederick T.
TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSER: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512th Carolina
COUNTRY: USA

ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rml

Page 10

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,564
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCE/DOCKET NUMBER: 5051-409
TELEPHONE: 919-854-1400
TELEFAX: 919-854-1401
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1657 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1548
US-08-948-564-11

```

Query Match 5.0%; Score 79.6; DB 3; Length 1657;
Best Local Similarity 49.1%; Pred. No. 4.3e-17;
Matches 276; Conservative 0; Mismatches 274; Indels 12; Gaps 2;

```

QY 956 ACATATATGCGCGCTTTGGAGATGATATTCAGAGGACAGACAGTGGCGTCTTA 1015
DB 914 ACATATAAGCGGTGTCATGAACCTTTATAGAGGAGACAGCCGAGTCCGACAA 973
QY 1016 TCGAGTGAATCCTCGCTAGATGTCCTTCATCCAGATATCAATCAACGCTACAAAACG 1075
DB 974 TAGTTTGGGCAATGAATGACATGTTGAGAAATCCCAATGATGAGCAAGTTCAAGAG 1033
QY 1076 AGCTGATCACTAGTATCGGGAATTAAGAGCCCTAGATGATCTGACTTGCTTCACTTC 1135
DB 1034 AAGTGAAGAACTTATTCGCTGACAAAGATTTCAATTAAGAAATGATGTCGAAAGCCTTC 1093
QY 1136 CATATCTAAGCGCTGTGTGAAGAAAGATTAAGGCTTCATCTCCAGGCCACTTCTAT 1195
DB 1094 CTATCTCAAGGACATGATGAGGACATTAAGATTAATCCACCTTCACTACT-- 1151
QY 1136 CATGGGCCCGTTGGCCATTAACAGACAGATGCTGATGCTGCTTCTTCCGAGAGGA 1255
DB 1152 -TTTCCCAAGGGTAACTGAATGAAATGCAATGCAATGAAAGGTACGAAATTCAGCCAAA 1210
QY 1256 CCACAGCAATGATGTAACATGATGAGGCGGTATGATGATCAACGATGAGGCTTATCCTT 1315
DB 1211 CTATGATGATTTAATGATGATGAGGACATTAAGCAAGGACCCCTGAAATTTGGGAAGCCTG 1270
QY 1316 TGGAGTTTAACCTTAAGAGTTCGTGGCAAAAGAGTGAAGGTTTCGGTCTTG 1375
DB 1271 AAAAAATTTTCCCCCAAGGTTCTTGAAGTTCATGAGATTAAGGGAATAT-- 1326
QY 1376 GGTGCGAATTTGAGACTTGACCTTCCGGTCCGGGTCTCGGATTTGGCCCGGGAAGATC 1435
DB 1327 -----GAGTTTAAGTATCCCGTTCCTTGGTCTGGAAGAGAAATGCTCTCGGAAGCA 1381
QY 1436 TTGGTTTAACTACCGTTATGTTTGAACGCGCATGATTAACATGATTTGAATGGGAC 1495
DB 1382 TGGGATATTTGATGATGAGCTTCTCTTGTCTAATCTATTCACACGTTTATGGGAAAG 1441
QY 1496 CGTCCGATGATTAACGCGCTTGA 1517
DB 1442 TGGCTTAAGGCTTCGACAAAGA 1463

```

RESULT 14
US-09-564-808-5
; Sequence 5, Application US/09564808
; Patent No. 6501004
; GENERAL INFORMATION:

```

APPLICANT: Selvaraj, Gopalan
APPLICANT: Nair, Ramesh B
APPLICANT: Joy IV, Richard W
APPLICANT: Keller, Wilfred A
APPLICANT: Datta, Raju S
TITLE OF INVENTION: Transgenic Reduction of Sinapine in Crucifera
FILE REFERENCE: 44618 Sequence Listing
Patent No. 6501004
CURRENT APPLICATION NUMBER: US/09/564,808
CURRENT FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: 60/132,800
PRIOR FILING DATE: 1999-05-06
PRIOR APPLICATION NUMBER: 2,270,417
PRIOR FILING DATE: 1999-05-06
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 1835
TYPE: DNA
ORGANISM: Brassica napus
FEATURE:
NAME/KEY: CDS
LOCATION: (2)..(1540)
US-09-564-808-5

```

Query Match 4.9%; Score 78.6; DB 4; Length 1835;
Best Local Similarity 48.7%; Pred. No. 1e-16;
Matches 284; Conservative 0; Mismatches 284; Indels 15; Gaps 2;

```

QY 925 CTCATGATGTCGATTAATTAATCCGACCCGACATTAATCCCGTCTTTGGAGATGATA 984
DB 872 CTTCAGAAATTCATCAATTAATCCGACATTAATCCGACATTAATCAATCAATGACCTCAAG 931
QY 985 TTCAGAGAAACAGACAGATTCGCTTCAATGAGATGATCCCTGATGATGATGATGATGAT 1044
DB 932 TTGGAAGAAACGAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 991
QY 1045 CATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1104
DB 992 AGCCCGAGAGATTAATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1051
QY 1105 GGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1164
DB 1052 CGGTGAGAAATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1111
QY 1165 TTGAGGCTTCAATCTTCAGAGCCCACTTCTATGATGAGGCGCTTGGCCATTAACAGACAG 1224
DB 1112 CTAAAGGTTACACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1165
QY 1225 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1284
DB 1166 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1225
QY 1285 TCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1344
DB 1226 GAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1285
QY 1345 AATGAGTGAAGTGAAGTTCGCTTCTTGGGTGAGATTTGAAGCTTCACTTCCGAG 1404
DB 1286 CCGGAGATGAGG-----ATTCAAGAGAAATGATGATGATGATGATGATGATGATGATGAT 1336
QY 1405 TCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1464
DB 1337 TCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1396
QY 1465 GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1507
DB 1397 GCTCATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1439

```

RESULT 15
US-09-091-432-1
; Sequence 1, Application US/09091432

```
/ Patent No. 5981837
/ GENERAL INFORMATION:
/ APPLICANT: Chapple, Clint
/ TITLE OF INVENTION: A Method For Regulation Of Plant Lignin Composition
/ FILE REFERENCE: 7024-325
/ CURRENT APPLICATION NUMBER: US/09/091,432
/ EARLIER FILING DATE: 1998-06-18
/ EARLIER APPLICATION NUMBER: PCT/US96/20094
/ EARLIER FILING DATE: 1996-12-19
/ EARLIER APPLICATION NUMBER: US 60/009,119
/ EARLIER FILING DATE: 1995-12-22
/ EARLIER APPLICATION NUMBER: US 60/013,388
/ EARLIER FILING DATE: 1996-03-14
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: Microsoft Word 2.0C
/ SEQ ID NO 1
/ LENGTH: 1838
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
US-09-091-432-1
```

```
Query Match      4.9%; Score 78.6; DB 2; Length 1838;
Best Local Similarity 48.7%; Pred. No. 1e-16;
Matches 284; Conservative 0; Mismatches 284; Indels 15; Gaps 2;

QY 925 CTCGATGCTTCAGATTAATTATCCGACCCGGACATATGCGCGTCTTTGGAGATGATA 984
DB 908 CTTCAAAATTCATCAACTTACCCGTGACAAATCAAGCAATCATGAGCGTTATG 967
QY 985 TTCAGAGAACAGACACAGTTGGGCTTATCGATGGATCCTCGTAGATGTCCTT 1044
DB 968 TTGAGAGAACGAAACGGAGTAGCGTGGGATAGAGTGGCCCTTAACGAGTTATTACGG 1027
QY 1045 CATCCAGATATGCAATCAACGATCAAAACGAGCTGATCAAGTATGCGGAAATCAGA 1104
DB 1028 AGCCCGAGAGATTAACCGGTCACAAAGAACTGCCGAAGTCGTTGACTTGACAGA 1087
QY 1105 GCCCTAGATGATCTGATCTGCTTCTTCATCTAATCAAGGCTGTGGTGAAGAAGTA 1164
DB 1088 CGAGTTGAAGAAATCCGACATCGAAGATGATCTTATCTCAATGCACTCAAGAAACC 1147
QY 1165 TTGAGCTTCATCTCCAGGCCCATCTTATCATGAGGCCGTTGGCCATTAACAGACAG 1224
DB 1148 CTAAGAGATGACCCACCGATCCCTCTCT-----CCTCCACGAAACCGCGAGGACACT 1201
QY 1225 ATGCTGATGATGCTGCTGTTGCGGACGAGGACCAAGCAATGATGTAACATGAGGCCGTA 1284
DB 1202 AGTATGACGGTTTCTTCATTTCCCAAGAAATCTCGTGTGATGATCAACGCGTTGGCATA 1261
QY 1285 TCGCATGATCCACAGCTGTGGGTTGATTCCTTTGAGTTTAAACCTGAGAGTTTGTGCA 1344
DB 1262 GGAACGCAACCACTTGTGACTGACCCGAGACAGCTTACCATCGAGGTTTGGAA 1321
QY 1345 AAGAGAGTGAAGTGAAGTTTGGCTTCTTGGGTCGAGTTTGAGACTTGACCTTTGGGG 1404
DB 1322 CCGGGCGTACCG-----ATTCAAGGAGGCAATTTGAGTTTATACCGTTTGGGG 1372
QY 1405 TCGGATGCTGGAATTTGCCCCGGGAAAGATCTTGTTTAACTACCGTTATGTTTGGACG 1464
DB 1373 TCGGGTGTGATGCTGCCCCGGGATGCACTAGGGTTATACGCGCTTAGCTTAGCGGTG 1432
QY 1465 GCCATGATGTTACATGAGTTTGAATGGGACCGTCCGATGATA 1507
DB 1433 GCTCATATATTAATGCTTCACGTGGAATTAACCTGATGGGA 1475
```

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

11088.430 Million cell updates/sec

AJ553770 *Arabidopsis*

AJ553770 *Arabidopsis*

ALIGNMENTS

RESULT 1
AC005819/c
LC005
AC005819
55870 bp
DNA
linear
PLN 11-MAR-2002

DEFINITION Arabidopsis thaliana chromosome 2 clone T3A4 map C1C06C03, complete sequence

ACCESSION
NUMBER

KEYWORDS
SOURCE
ORGANISM

KEYWORDS
SOURCE
ORGANISM

GenBank: *
1: gb_ba: *

REFERENCE
1 (bases 1 to 55870) "1.8 kb EcoRI-ClaI fragment
rosids; eurosid II; brassicales; brassicaceae; arabiopsys.

[illegible]

```

EILRSPOGKKTGGSGTHSSRSVSHGTGKGRRLQSSSSSSDEGKKTKTKVNVKKN
SNLPMKAFMNSNVQMINNSIVNSTASHHDPGVHLTKISRKPGSDNGFHVSDYGNNGY
TN"
repeat_region
723. .810
/rpt_family=" (CACCAT)n"
repeat_region
1345. .1382
/rpt_family="AT_rich"
repeat_region
complement (1904. .1925)
/rpt_family="AT_rich"
3740. .5558
/gene="At2g46640"
/note="synonym: T3A4.2; predicted by genscan and
genefinder"
join (<3740. .4072,4525. .4839,5083. .5215,5407. .>5558)
/gene="At2g46640"
join (3740. .4072,4525. .4839,5083. .5215,5407. .5558)
/gene="At2g46640"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC69921.1"
/db_xref="GI:3831438"
/translation="MWKEIDGVKRNKKKRNKNEGISEIKTKAILDQGLVADLWPF
DGVLTIGTGFDTLNFKEEDMDDECGSDLLVYIDGSIINVDQESDPLISVKNK
FYDHEVDGADTDHFGSIKTAETPAVTAALAEIPEKRTTLAELEFMBHDGDTYD
HCKKKNPNLDGEVEKYKNGSKLSKRFSEFKKLQVMSKKEENDLPFKKRCMI
KRMKKKHPDVKATKAKDVPYKPTRNCEALSHYLKIQCGVGIWMTVEWYCYLL
TNIKCLVFPVWHIVFYVVICLTITYNASP"
6216. .7156
/gene="At2g46650"
/note="synonym: T3A4.3; supported by full length cDNA:
Ceres:3743"
join (6216. .6447,6552. .6618,6704. .7156)
/gene="At2g46650"
join (6370. .6447,6552. .6618,6704. .6957)
/gene="At2g46650"
/codon_start=1
/product="putative cytochrome b5"
/protein_id="AAC69922.1"
/db_xref="GI:3831439"
/note="synonym: T3A4.4"
complement (join (<8165. .8782,9006. .>9980))
/gene="At2g46660"
complement (join (8165. .8782,9006. .9980))
/gene="At2g46660"
/codon_start=1
/product="putative cytochrome P450"
/protein_id="AAC69923.1"
/db_xref="GI:3831440"
/translation="MATKLBSLIFALSKCSVLQNTLAFSLAVTIIWLAISLFLW
TYPGPAWKGKYLFGRLISGYKGVNIPGKGFPIVSGMSLMSSTLARRJADAAERF
GAKRMAFSIGETRIYVINCNPVAKELINSPEVADRPKESAGSLMEFRAIGFAPHV
VYRTLRRTASNLHSTKOIRRAETORRIVISSOMEFLEKSSNBPCEYRELTKASIN
NMKCVPGQEVTLKQHYELPEMDEGVILGTMWTHLPMLSBPDRRLRSKSTL
VKVNRVSRITISEHRNOTGDIPEPVDVLSHSGDLSLSPDIATVEMIFRGDT
VAVILWILARVLRHDMOSTVQNSLSDVVGSKRLDSDSLASPLITRAVKEVLRH
PEGPLISVARLAIPTDITVGRVPAQTAMNMAVSDPVMWVDPLEFKERVAKE
GEVERSVLGSDLRALTPFGSGRIICPGKMLGFTTVVFWTAMMLHEFWGSPDNGVDLS
EKLRLSCMANPLPAKLRRRS"
10124. .10180
repeat_region
/rpt_family="AT_rich"
complement (14979. .15002)
/rpt_family="(A)n"

```

```

repeat_region
15546. .15568
/rpt_family="AT_rich"
repeat_region
17214. .17366
/rpt_family="AT_rich"
repeat_region
17478. .17503
/rpt_family="(A)n"
repeat_region
complement (17523. .17543)
/rpt_family="AT_rich"
repeat_region
complement (17647. .17688)
/rpt_family="AT_rich"
repeat_region
complement (17874. .17894)
/rpt_family="AT_rich"
repeat_region
complement (18001. .18060)
/rpt_family="AT_rich"
repeat_region
complement (19152. .19796)
/gene="At2g46670"
/note="synonym: T3A4.5; predicted by genscan and
genefinder"
complement (join (<19152. .19310,19344. .>19796))
/gene="At2g46670"
complement (join (19152. .19310,19344. .19796))
/gene="At2g46670"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC69924.1"
/db_xref="GI:3831441"
/translation="WYLTGCFRFEBSKSAKAVVALESTGSEPKTFTSHKXLRVSD
QSSSTTSNQNIGSSSVSFNQVLCSTVYQKQDSPIPVSNREKASKEVEGSGS
TNEGIKAGSSSTKKEBSAKORMSORAEALMRKRCDFPKHLKQIQDOT
DSMEWYKGLQVRYSRKKLAQRPRVKGQVFRVNSDASTKS"
19989. .21343
/note="molecular marker Atcb7"
complement (20037. .21512)
/gene="At2g46680"
/note="synonym: T3A4.6; identical to SP:146897; supported
by cDNA: gi 15027938.gb AY045826.1"
complement (join (20037. .20039,20111. .20756,20977. .>21512))
/gene="At2g46680"
complement (join (20340. .20756,20977. .21361))
/gene="At2g46680"
/codon_start=1
/product="homeodomain transcription factor (ATHB-7)"
/protein_id="AAC69925.1"
/db_xref="GI:3831442"
/translation="WTGGEYSAPMASSAPPLTMKKMKSNHNNKORRPFDEQIKSL
EMWFSETRLEPRKRVQIAREIGLOPROVAILMFKRABWKSQLEETVNLIRQVND
IASQFESLKEKQALVSELQRIKATQKQBEERQCGDDAVVALSSTHSEBREN
RRKREVARPEMEMKDDKHGHCWCDHHDYEDDDGYSNNIRRETFGGFEEBPHLMN
IVEPADSCITSSDDWRFKSDPTTLTLDSSNNYPWRDWS"
complement (22636. .22664)
/rpt_family="AT_rich"
23532. .23571
/rpt_family="AT_rich"
repeat_region
23760. .23783
/rpt_family="AT_rich"
repeat_region
24595. .24656
/rpt_family="(CAAAA)n"
complement (25135. .25155)
/rpt_family="AT_rich"
25903. .25923
/rpt_family="AT_rich"
repeat_region
complement (26667. .26771)
/rpt_family="AT_rich"
repeat_region
complement (26830. .26901)
/rpt_family="Rf:ATREP1"
ATREP1 ATREP1 dispersed repeat from
GB1D6608 - a consensus.
30015. .30038
/rpt_family="(TAAA)n"
complement (31911. .31933)
/rpt_family="AT_rich"
complement (32444. .32466)
/rpt_family="AT_rich"
35276. .35832

```



```

gene      /rpt_family="AT-rich"
           /gene="At2g4620"
           /note="synonym: F13A10.5"
           /gene="At2g46520"
           /gene="At2g46520"
           /gene="At2g46520"
           /product="putative cellular apoptosis susceptibility
           protein"
           /protein_id="AAD20163.1"
           /db_xref="GI:4415933"
           /rpt_family="MEMBER:VPLSOCIANTLSP1PPEPPTAERALSADADQANGL
           AVRLVAEPAIDBQTRHAAVNFKXHLBRKMTPADDSQISPIVDEKQOITLIVSLM
           LSAPRIQSLBALVIGADHPFKAMPRLPELANTONMALADYISVNGILGTAS
           S1FKFSYRTDALFVLDKCLDNFAALTEIFLKTSLIDSAASGSPILPELF
           ESQRLCCTIFYSINFDLPEDHKEWGEFKYLSNYPALSTEGTLVDLDR
           AAIENINHYIEKNEBEFQGPLNEFASVVTLLRDVSGSPSDQATTAIKFLTSVT
           SVHHLFAGDNVKEICQSIIVPNVSLRDEBEIPEMNVIPIRIDMGSDYDERRI
           ACHLKGATANKYQWTEYVSLIEIOLISFSPANSMKDKCALIVLSLSTKAG
           GASVSTDLIDVQFFANITLPELQSDVNSFPMKAGSLKPTIMRSHIPPEMNOIF
           PELVRLKAEISVHSYASCIKELIVBEGARKGRVAGDLSFPLQNTNFDAL
           KEPSENOYLMKCIWVAVADISAEVAGPCIGLITLSBVCNPKNP1FNH1FE
           SVAIVRACERDISLISAFETSLPSLOMLIANDITEFLPYGQLAQVELNRPFL
           SPYMOIFILLISPSMKSGNVPALVRLLOAPLOKAPREVOENRLOQVIGIPEKLV
           ASPSDQGFYITNTIENDYSVIAPYKVGSALEFTRQNKTKVFKQKSLVFMSL
           PLYKHGOALVETMNTVOPNIIITAIHEHFMILKIMSGMEVKTLLVAVALICETP
           ALDPSAKLKMGMKMDISTIVYSRPGESEVLDPEMPBISSENVGTAAAFVRLHMGKX
           EBDPLDIOPKQFLVASVSLSSASPGKIPQ1IGENLEQAOQTALIQLCARVNGCIA
           "
repeat_region 25088..25183
               /rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.chaliana
               microsatellite repeat region (clone 74A) #Rf:ATR0046|X93607
               74A"
repeat_region 25090..25183
               /rpt_family="GA"n"
               /rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.chaliana
               microsatellite repeat region (clone 74A) #Rf:ATR0046|X93607
               74A"
repeat_region 25090..25166
               /rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.chaliana
               microsatellite repeat region (clone 74A) #Rf:ATR0046|X93607
               74A"
repeat_region 25090..25162
               /rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.chaliana
               microsatellite repeat region (clone 74A) #Rf:ATR0046|X93607
               74A"
repeat_region 25090..25158
               /rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.chaliana
               microsatellite repeat region (clone 74A) #Rf:ATR0046|X93607
               74A"
repeat_region 25090..25154
               /rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.chaliana
               microsatellite repeat region (clone 74A) #Rf:ATR0046|X93607
               74A"
Query Match 100.0%; Score 2009; DB 8; Length 84825;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2009; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 241 CTCCTTCCTGCGCGTCACAATCATCTGCTCGCATATCTCTTCTTATGACCTATCC 300
DB 79978 CTCCTTCCTGCGCGTCACAATCATCTGCTCGCATATCTCTTCTTATGACCTATCC 79919
QY 301 CGGTGACCTGCTTGGGGGAAATACCTCTTGGCGGCTTAATATCCGGTTCATCAAAAC 360
DB 79918 CGGTGACCTGCTTGGGGGAAATACCTCTTGGCGGCTTAATATCCGGTTCATCAAAAC 79859
QY 361 CGAAAGCTATATCCGGGTCAAAAGGCTCCCTTGTGTGAAGCATGCTCATGTC 420
DB 79858 CGAAAGCTATATCCGGGTCAAAAGGCTCCCTTGTGTGAAGCATGCTCATGTC 79799
QY 421 AAGCACTCTGACCTACCGAGCAATCGCTGATGCAAGTATGGAGCCCAAGGCT 480
DB 79798 AAGCACTCTGACCTACCGAGCAATCGCTGATGCAAGTATGGAGCCCAAGGCT 79735
QY 481 CATGCTTTGACCTTAGAGAGAGACTCGCGGATCGTCACTGCAATCCGACGTAGGAA 540
DB 79738 CATGCTTTGACCTTAGAGAGAGACTCGCGGATCGTCACTGCAATCCGACGTAGGAA 79679
QY 541 AGAATTCGAAATAGCCCGCTTTTGTCTGATCGACCGGTTAAAGATCGCTTACTACT 600
DB 79678 AGAATTCGAAATAGCCCGCTTTTGTCTGATCGACCGGTTAAAGATCGCTTACTACT 79619
QY 601 GATGTTTAAAGAGCAATTTGGTTTGAACAACCGGTTTACTGGGGAACGCTTGGCG 660
DB 79618 GATGTTTAAAGAGCAATTTGGTTTGAACAACCGGTTTACTGGGGAACGCTTGGCG 79559
QY 79618 GATGTTTAAAGAGCAATTTGGTTTGAACAACCGGTTTACTGGGGAACGCTTGGCG 79559
QY 661 TATGCTTCGAACCATCTCTTATAGTACAAAACAATCAGAAAGCCGAGACGAGACG 720
DB 79558 TATGCTTCGAACCATCTCTTATAGTACAAAACAATCAGAAAGCCGAGACGAGACG 79499
QY 721 AGTATCTCAAGCCAGATGTTGATGCTTCTTGAAGAAACAGATGATACGAAACCCGTT 780
DB 79498 AGTATCTCAAGCCAGATGTTGATGCTTCTTGAAGAAACAGATGATACGAAACCCGTT 79439
QY 781 TGTTCGATGATGCTTAAACCGCGCTGCTTAAACAATGATGCTCTGATTCGACA 840
DB 79438 TGTTCGATGATGCTTAAACCGCGCTGCTTAAACAATGATGCTCTGATTCGACA 79379
QY 841 AGAGTATGAGCTTGAAGAAACCATGTTAGTTACGTGAATGGTCAAGAGCTTATGA 900
DB 79378 AGAGTATGAGCTTGAAGAAACCATGTTAGTTACGTGAATGGTCAAGAGCTTATGA 79319
QY 901 TTTGCTCGAAGCTTGAATTTGAGCTGATCACTCTTGGCTACCGAGTTTACTCTCA 960
DB 79318 TTTGCTCGAAGCTTGAATTTGAGCTGATCACTCTTGGCTACCGAGTTTACTCTCA 79259
QY 961 AAGACTCGGCTCTGATGTTCCACACTGTCACCAAGAGTAAACGGTTGTATCCGGAT 1020
DB 79258 AAGACTCGGCTCTGATGTTCCACACTGTCACCAAGAGTAAACGGTTGTATCCGGAT 79199
QY 1021 TATATCCGAACACGCTATCAAAACCGGATTTGGCTCGCGAATTTGGCTGCGATTTGCT 1080
DB 79198 TATATCCGAACACGCTATCAAAACCGGATTTGGCTCGCGAATTTGGCTGCGATTTGCT 79139
QY 1081 CTCCTTCATGTTGATGATTAATATCCGACCCGAGCAATAATGCGCTTCTTGGGTATG 1140
DB 79138 CTCCTTCATGTTGATGATTAATATCCGACCCGAGCAATAATGCGCTTCTTGGGTATG 79079
QY 1141 CACACCATTTATTTGATTAATTTCTTAATTAATTTGTTGAAATTTGCTTAGGATAT 1200
DB 79078 CACACCATTTATTTGATTAATTTCTTAATTAATTTGTTGAAATTTGCTTAGGATAT 79019
QY 1201 TTAGATTAACAATGAATTTGAGACTCAATGTCAGTGTGTGGAATTAATTAAGACT 1260
DB 79018 TTAGATTAACAATGAATTTGAGACTCAATGTCAGTGTGTGGAATTAATTAAGACT 78959
QY 1261 AGAAGTTTTTGTGGAATCAATTAATTTGATTAATTTTAAATTTTAAAGTTTCTATAA 1320
DB 78958 AGAAGTTTTTGTGGAATCAATTAATTTGATTAATTTTAAATTTTAAAGTTTCTATAA 78899
QY 1321 AAGTAAAGATCAATATTTTGTAGTTATTAATTAACAGAGATGATTTACAGAGACAG 1380
DB 1380 AAGTAAAGATCAATATTTTGTAGTTATTAATTAACAGAGATGATTTACAGAGACAG 1380

```

Db		78698	AAGTAAAAATGTACTAAATAATTTTGAGTATTATTCACGAGATGAATTCACAGGAACAG	78839
Qy		1381	ACACAGTTGGCGGTCTTAATGAGTGCATCCTCCCTAGATGAGTGCCTTCATCCAGATATGC	1440
Db		78638	ACACAGTTGGCGGTCTTAATGAGATGATCCTCCCTAGATGAGTGCCTTCATCCAGATATGC	78779
Qy		1441	AATCAACGGTACAAAAGACCTGATCAAATAGTAGTCGGGAAATCAAGGCCCTGATGAT	1500
Db		78778	AATCAACGGTACAAAAGACTGATCAATGATATCCGGAAATCAAGGCCCTGATGAT	78719
Qy		1501	CTGACTGGCTTCACCTTCCAATATCTAACCGGCTGTGTTGAAGAAGATTGAGGCTTCATC	1560
Db		78718	CTGACTGGCTTCACCTTCCAATATCTAACCGGCTGTGTTGAAGAAGATTGAGGCTTCATC	78659
Qy		1561	CTCCAGGCCCATCTTCATCATGAGGCCGCTTTGGSCATTAACAACAAGATCGTTGATGTC	1620
Db		78658	CTCCAGGCCCATCTTCATCATGAGGCCGCTTTGGSCATTAACAACAAGATCGTTGATGTC	78539
Qy		1621	GTCCTGTTCCGCAAGGACACAACAGCAATGCTGAACATGTGGGCCGTATCCGATATCAC	1680
Db		78598	GTCCTGTTCCGCAAGGACACAACAGCAATGCTGAACATGTGGGCCGTATCCGATATCAC	78539
Qy		1681	ACGTTGGGTTGATACCTTTGGAATTTAAACTGAGAGTTTCGGGCAAAAGAAGTGAGG	1740
Db		78538	ACGTTGGGTTGATACCTTTGGAATTTAAACTGAGAGTTTCGGGCAAAAGAAGTGAGG	78479
Qy		1741	TGAGTTTTGCGGTTCTTGSGTCGGAATTTGAGACTTGACACCTTCGGGTCGTCCGA	1800
Db		78478	TGAGTTTTGCGGTTCTTGSGTCGGAATTTGAGACTTGACACCTTCGGGTCGTCCGA	78419
Qy		1801	TTTGCCCCCGGGAAGAACTCTGTGTTTAACTACACCTTAATGTTTGACACGAGATGATAC	1860
Db		78418	TTTGCCCCCGGGAAGAACTCTGTGTTTAACTACACCTTAATGTTTGACACGAGATGATAC	78359
Qy		1861	ATGAGTTTGAATGGGGAACGTCCTCGATGTAACGGCGTTGACTTATCTGAGAAACTGAGGC	1920
Db		78358	ATGAGTTTGAATGGGGAACGTCCTCGATGTAACGGCGTTGACTTATCTGAGAAACTGAGGC	78299
Qy		1921	TTTCTTCGAGATGCGCTAATCCTCTTCCTGCTAAATTTGGCCGTAGGCCAGTTAAAAA	1980
Db		78298	TTTCTTCGAGATGCGCTAATCCTCTTCCTGCTAAATTTGGCCGTAGGCCAGTTAAAAA	78259
Qy		1981	AAGAGCTCATATGAGCAATTAAGAGATTTT	2009
Db		78238	AAGAGCTCATATGAGCAATTAAGAGATTTT	78210
RESULT 3				
ATH552847/c				
LOCUS			740 bp DNA linear pin 29-MAR-2003	
DEFINITION			Arabidopsis thaliana T-DNA flanking sequence, left border, clone 34EB11.	
ACCESSION			AU552847	
VERSION			AU552847.1 GI:29368998	
KEYWORDS			left border; T-DNA flanking sequence.	
SOURCE			Arabidopsis thaliana (thale cress)	
ORGANISM			Arabidopsis thaliana	
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.				
REFERENCE			1	
AUTHORS			Brunaud, V., Balzerquc, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Crnau, C., Derose, R., Pelletier, G., Lepoint, L., Caboche, M., and Lecharny, A.	
TITLE			T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites	
JOURNAL			EMBO Rep. 3 (12), 1152-1157 (2002)	
MEDLINE			22363535	
PUBMED			12446565	
REFERENCE			2 (bases 1 to 740)	
AUTHORS			Balzerquc, S.	
TITLE			Direct Submission	

COMMENT	FEATURES	BASE COUNT	ORIGIN
JOURNAL	Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE		
PCRs were performed on DNA from transformants of <i>Arabidopsis thaliana</i> plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at http://dbsgap.versailles.inra.fr/publiclines/ . This sequence has been generated in the framework of the French plant genomes program 'Genoplante' (http://www.genoplante.com and http://genoplante-info.inbiohogen.fr).			
source	Location/Qualifiers		
	1. 740		
	/organism="Arabidopsis thaliana"		
	/mol_type="genomic DNA"		
	/cultivar="Massilllewkija"		
	/db_xref="taxon:3702"		
	/clone="345B11"		
	/clone_id="Arabidopsis thaliana T-DNA insertion lines"		
	1. 740		
	/note="T-DNA flanking sequence"		
	left border"		
Query Match	11.5%; Score 232, DB 8; Length 740;		
Best Local Similarity	99.4%; Pred. No. 2.2e-117;		
Matches 332; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		
QY	360 CCGGAAACGTTATCCCGGATCCAAAGGCTCCCTTGGTTGGAAGCATGTCATCATGT	192 a	172 c
DB	429 CCGGAAACGTTATCCCGGATCCAAAGGCTCCCTTGGTTGGAAGCATGTCATCATGT	197 g	179 t
QY	420 CAAGCACTTACGCTACCGACGAATCGCTGATGACGTGAAATTCGAGCCAAAGGC		
DB	369 CAAGCACTTACGCTACCGACGAATCGCTGATGACGTGAAATTCGAGCCAAAGGC		
QY	480 TCATGCGCTTCAGGTTAGAGAGAGAGCTGCGTGAATCGCATCGTGAATCCCGAGTACGA		
DB	309 TCATGCGCTTCAGGTTAGAGAGAGAGCTGCGTGAATCGCATCGTGAATCCCGAGTACGA		
QY	540 AAGAGATTCGGAATAGCCCGGTTTTCGTAATGACCGGTTAAAGATCGGCTTACTAC		
DB	249 AAGAGATTCGGAATAGCCCGGTTTTCGTAATGACCGGTTAAAGATCGGCTTACTAC		
QY	600 TGATGTTAAACAGAGCAATGATTTTGCACACACGCGTGTATCTGGCGAAGCTTGCC		
DB	189 TGATGTTAAACAGAGCAATGATTTTGCACACACGCGTGTATCTGGCGAAGCTTGCC		
QY	660 GTATCGCTTCGAACCATCTCTTTGATCAAAACA		
DB	129 GTATCGCTTCGAACCATCTCTTTGATCAAAACA		
RESULT 4			
LOCUS	ATH553770		
DEFINITION	Arabidopsis thaliana T-DNA flanking sequence, left border, clone 371B10.		
ACCESSION	AJ553770		
VERSION	AJ553770.1		
KEYWORDS	left border; T-DNA flanking sequence.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Brumaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Craud, C., Dehose, R., Pelletier, G., Lepoint, L., Caboche, M., and Leclercq, A.		

TITLE T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites
 JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
 MEDLINE 22363535
 PUBMED 12446565
 REFERENCE 2 (bases 1 to 504)
 AUTHORS Balzerque, S.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE

COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbsgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).
 Location/Qualifiers

FEATURES
 source 1..504
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="MassiliaeskiJa"
 /db_xref="taxon:3702"
 /clone="371B10"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"

misc_feature 1..504
 /note="T-DNA flanking sequence
 left border"

BASE COUNT 180 a 128 c 71 g 125 t
 ORIGIN

Query Match 7.0%; Score 141; DB 8; Length 504;
 Best Local Similarity 99.5%; Pred. No. 1.7e-66;
 Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1714 AGAGTTCTGTCGCAAAAGAGTGAAGTTTCGTTCTTGAGTTCGATTGAGAC 1773
 Db 497 AGAGTTCTGTCGCAAAAGAGTGAAGTTTCGTTCTTGAGTTCGATTGAGAC 438
 QY 1774 TTGACCTTTGGGTCGGGTCGTGGATTGCCCCGGAGAATCTTGGTTTACTACCG 1833
 Db 437 TTGACCTTTGGGTCGGGTCGTGGATTGCCCCGGAGAATCTTGGTTTACTACCG 378
 QY 1834 TTATGTTTGGACGGCGATGATGTTCATGAGTTTGAATGGGACCGTCCGATGTTACG 1893
 Db 377 TTATGTTTGGACGGCGATGATGTTCATGAGTTTGAATGGGACCGTCCGATGTTACG 318
 QY 1894 GCGTTGACTTAT 1905
 Db 317 GCGTTGACTTAT 306

Search completed: January 16, 2004, 18:05:05
 Job time : 7412 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 09:30:00 ; Search time 4825 Seconds

(without alignments)
10119.729 Million cell updates/sec

Title: US-10-022-025A-1
Perfect score: 2009
Sequence: 1 aataataataataatgta.....atagagaattagatttc 2009

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estlum:*
16: em_estlom:*
17: em_ges_hum:*
18: em_ges_huv:*
19: em_ges_pln:*
20: em_ges_vrt:*
21: em_ges_fun:*
22: em_ges_nam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rtd:*
26: em_ges_phg:*
27: em_ges_vrl:*
28: gb_ges1:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	566.4	28.2	712	29	BZ428523 BONHW74TF
2	550.6	27.4	784	28	BH492184 BOHFJ050TR
3	502.6	25.0	809	29	BZ428529 BONHW74TR
4	479	23.8	674	28	BH451907 BOGSO44TR

Result No.	Score	Query Match	Length	DB ID	Description
5	472	23.5	804	28	BH431683 BOHR746TR
6	458.8	22.8	745	28	BH510149 BOGTQ49TR
7	421.6	21.0	727	28	BH527239 BOHBQ06TR
8	419.6	20.9	659	29	BZ466758 BONND11TR
9	396.6	19.7	776	29	BZ443616 BONFE28TR
10	382.6	19.0	576	28	BH451899 BOGSO44TF
11	337.6	16.8	705	28	BH77803 Cde14a11.
12	330.8	16.5	732	28	BH534160 BOHCZ39TF
13	311.2	15.5	826	14	CB344213 CA48EN000
14	304	15.1	638	14	CB343993 CA48EN000
15	289.2	14.4	755	28	BH581933 BOHRM09TR
16	288.6	14.4	612	9	AV826713 AV826713
17	283.4	14.1	764	28	BH652028 BOHRX14TF
18	279.8	13.9	791	12	BG839564 CM01_12H0
19	271.2	13.5	455	9	A1995549 701674596
20	254.6	12.7	846	13	BQ518647 EST626062
21	252	12.5	549	28	BH250286 BOGAJ62TR
22	251.8	12.5	549	12	B1973185 sal5a11.
23	238.8	11.9	691	9	AM132420 se04f07.Y
24	238.4	11.9	740	14	CB344214 CA48EN000
25	236.6	11.8	540	29	BZ436712 BOAS69TR
26	233.6	11.6	739	28	BH547739 BOGXW70TF
27	233.2	11.6	489	10	BE346914 sp32f12.Y
28	231.2	11.5	641	12	BM063648 KS01057H1
29	224.4	11.2	835	10	BG447296 GA_Eb004
30	220.4	11.0	454	10	BE210324 s042d01.Y
31	217.6	10.8	810	10	BG583985 EST485745
32	213.2	10.6	537	9	AL365677 MCB401G10
33	212.6	10.6	468	9	AM586022 EST317645
34	212	10.6	653	10	BF595944 su67h07.Y
35	210.6	10.5	764	13	BQ403090 GA_Ed005
36	208.6	10.4	531	10	BE131362 sm71e11.Y
37	207.4	10.3	524	10	BE398607 EST430330
38	204.6	10.2	720	10	BG468529 EST510148
39	201.6	10.0	616	14	CA030582 HX07H18T
40	199.6	9.9	262	29	BZ380056 SALK_1145
41	198.8	9.9	659	13	BU820409 UAS4DPH06
42	196.8	9.8	773	28	BH013345 TDGAW84TH
43	193.2	9.6	888	29	CC330440 OGIBD13TV
44	189.6	9.4	837	14	CB657900 OSUNE131
45	189	9.4	807	29	CC336997 OGRAB60TV

ALIGNMENTS

RESULT 1
LOCUS BZ428523
DEFINITION BONHW74TF BO.1.6.2.KB lot Brassica oleracea genomic clone BONHW74.
GENOMIC SURVEY SEQUENCE.
ACCESSION BZ428523
VERSION BZ428523.1 GI:26671493
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids / euroside II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 712)
Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other GSSs: BONHW74TR
COMMENT Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: 1F
Class: sheared ends.

Tue Jan 20 17:43:19 2004

us-10-022-025a-1.rst

Page 2

FEATURES

Source

Location/Qualifiers

1..712
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BONHW74"
/clone_1lb="BO_1.6.2 KB tot"
/note="Vector: pHOSt; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into pHOSt using BstXI linkers"

BASE COUNT

168 a 179 c 177 g 188 t

ORIGIN

Query Match 28.2%; Score 566.4; DB 29; Length 712;
Best Local Similarity 87.2%; Pred. No. 1.2e-102;
Matches 621; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

411 CACTCATGTCAAGAGCTCTGAGCTCAGCGAGATCGCTGATGCGAGTAATTGGAG 470
1 CTCCTCATGTCCAGAGCTCTGAGCTCTCCGCGAATCGCAGCTCTGCGAGAGATTGGAG 60

471 CCAAGAGGCTTATGCTTACGCTTACGAGAGAGCTCGGATCGTCAAGTCCCG 530
61 CCAAGCGCTTTAGCTTACGCTTACGAGAGAGAGCTCGGATCGTCAAGTCCCG 120

531 ACGTAGCGAAGAGATTCTGATAGCCGGTCTTTGCTGATCGACCGGTTAAGATCGG 590
121 ACGTTGCTAAGAGATTCTGATAGCCGGTCTTTGCTGATCGACCGGTTAAGATCGG 180

591 CTTACTCATGTATTTTACAGAGCAATGGTTTGGACACAGCGGTTTACCGGAGAA 650
181 CTTACTCATGTATTTTACAGAGCAATGGTTTGGACACAGCGGTTTACCGGAGAA 240

651 CGCTTCGCGGATGCTTGAACCACTCTTTAGTACAAACCAATCAGAAAGACCGAGA 710
241 CTCGCGGCGATGCTTGAACCACTCTTTAGTACAAACCAATCAGAAAGACCGAGA 300

711 CGCAACGAGAGATCTCAAGCCAGAGAGTGTGATTTCTTGAAAAACAGATGTAGAG 770
301 CACAGAGCGAGATGCTCAAGCCAGAGAGTGTGATTTCTTGAAAAACAGATGTAGAG 360

771 AACCTGTTTGTGCTGATGCTTAAACGCGCTCGCTTACAGCATGATGTCTCTG 830
361 GCGTGTGTTTGTGCTGATGCTTAAACGCGCTCGCTTACAGCATGATGTCTCTG 420

831 TATTCGAGCAAGATGATGAGCTTGAAAAAACATGTTGATGATGATGATGATGATG 890
421 TTTTCGAGCAAGATGATGAGCTTGAAAAAACATGTTGATGATGATGATGATGATG 480

891 AAGGTTAGATTTGCTCGGAGAGTGAATGGAATGGAATGGAATGGAATGGAATGGA 950
481 AAGGTTAGATTTGCTCGGAGAGTGAATGGAATGGAATGGAATGGAATGGAATGGA 540

951 TTGATCTCTCAAGAGCTCGGCTGATGCTTCAACTGTGACCAAGGTAACCGGTTTG 1010
541 TTGATCTCTCAAGAGCTCGGCTGATGCTTCAACTGTGACCAAGGTAACCGGTTTG 600

1011 TATCCCGAGTTATATCCGAGACCGGTAATCAACCGGTAATGCTGCTGATTTGCTCG 1070
601 TATCCCGAGTTATATCCGAGACCGGTAATCAACCGGTAATGCTGCTGATTTGCTCG 660

1071 ACGTTTGTCTCTCCCTCCATGCTTCAAGTAAATTAATCCGACCGGACATTAAT 1122
661 ACGTTTGTCTCTCCCTCCATGCTTCAAGTAAATTAATCCGACCGGACATTAAT 712

RESULT 2
BH492184 784 bp DNA linear GSS 13-DEC-2001

LOCUS BOHFJ05TR BOHF Brassica oleracea genomic clone BOHFJ05, genomic

DEFINITION survey sequence.

ACCESSION BH492184

VERSION BH492184.1 GI:17700288

KEYWORDS

GSS.
Brassica oleracea

SOURCE

Brassica oleracea

ORGANISM

Brassica oleracea

REFERENCE

1 (bases 1 to 784)

AUTHORS

Town, C.D., Van Aken, S., Utecherback, T., Koo, H. and Fraser, C.M.

TITLE

Whole genome shotgun sequencing of Brassica oleracea

JOURNAL

Unpublished

COMMENT

Other_GSSs: BOHFJ05TR

CONTACT

Contact: Chris Town

ORIGIN

TIGR

BASE COUNT

204 a 189 c 211 g 180 t

Query Match 27.4%; Score 550.6; DB 28; Length 784;
Best Local Similarity 83.8%; Pred. No. 1.7e-99;
Matches 651; Conservative 0; Mismatches 114; Indels 12; Gaps 2;

123 TCCT 182
774 TCT 715

183 CCTTAATCTTTGCTTCT 242
714 CCTTAATCTTTGCTTCT 655

243 CCTTAATCTTTGCTTCT 302
654 CCTTAATCTTTGCTTCT 595

654 CCTTAATCTTTGCTTCT 595
303 CCTTAATCTTTGCTTCT 362

303 CCTTAATCTTTGCTTCT 362
594 CCTTAATCTTTGCTTCT 544

594 CCTTAATCTTTGCTTCT 544
363 CCTTAATCTTTGCTTCT 422

363 CCTTAATCTTTGCTTCT 422
543 CCTTAATCTTTGCTTCT 484

543 CCTTAATCTTTGCTTCT 484
423 CCTTAATCTTTGCTTCT 542

423 CCTTAATCTTTGCTTCT 542
483 CCTTAATCTTTGCTTCT 542

483 CCTTAATCTTTGCTTCT 542
423 CCTTAATCTTTGCTTCT 542

423 CCTTAATCTTTGCTTCT 542
483 CCTTAATCTTTGCTTCT 542

483 CCTTAATCTTTGCTTCT 542
423 CCTTAATCTTTGCTTCT 542

423 CCTTAATCTTTGCTTCT 542
483 CCTTAATCTTTGCTTCT 542

483 CCTTAATCTTTGCTTCT 542
423 CCTTAATCTTTGCTTCT 542

423 CCTTAATCTTTGCTTCT 542
483 CCTTAATCTTTGCTTCT 542

483 CCTTAATCTTTGCTTCT 542
423 CCTTAATCTTTGCTTCT 542

423 CCTTAATCTTTGCTTCT 542
483 CCTTAATCTTTGCTTCT 542

483 CCTTAATCTTTGCTTCT 542
423 CCTTAATCTTTGCTTCT 542

423 CCTTAATCTTTGCTTCT 542
483 CCTTAATCTTTGCTTCT 542

483 CCTTAATCTTTGCTTCT 542
423 CCTTAATCTTTGCTTCT 542

423 CCTTAATCTTTGCTTCT 542
483 CCTTAATCTTTGCTTCT 542

483 CCTTAATCTTTGCTTCT 542
423 CCTTAATCTTTGCTTCT 542

423 CCTTAATCTTTGCTTCT 542
483 CCTTAATCTTTGCTTCT 542

483 CCTTAATCTTTGCTTCT 542
423 CCTTAATCTTTGCTTCT 542

423 CCTTAATCTTTGCTTCT 542
483 CCTTAATCTTTGCTTCT 542

QY 663 TCCTTCGAACCACTCTTTAGTACAAACAAATCAGAGAGCCGAGCCGACGAG 722
 DB 243 TCCTTCGAACCACTCTTCAGTCTTAAACAAATCAACGCGGAGAACGAGGCGAG 184
 QY 723 TGAATTCAGCCAGATGCTGAGTTCTTGAAGAAACAG---AGTACTTAACGAACTCTTT 779
 DB 183 TGAATTCAGCCAGATGCTGAGTTCTTGAAGAAACAGAGAGTACTTAACGAGGCTCTGTT 124
 QY 780 TTGTTCTGAGTGTCTTAAACGCGCTGCTTAAACAGATGCTCTGTAATTTGAGC 839
 DB 123 TTGTTCTGAGTGTCTTAAACGCGCTGCTTAAACAGATGCTCTGTAATTTGAGC 64
 QY 840 AAGAGTATGAGCTTTGAAAAAACATGTTAGTACTGTAATGCTGGAAGAGTT 896
 DB 63 AAGAGTATGAGCTTTGAAAAAACATGTTAGTACTGTAATGCTGGAAGAGTT 7

RESULT 3
 B2428529 809 bp DNA linear GSS 13-DEC-2002
 LOCUS B2428529 BO_1.6_2_KB_tot Brassica oleracea genomic clone BONHW74,
 DEFINITION genomic survey sequence.
 ACCESSION B2428529
 VERSION B2428529.1 GI:26671510
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 809)
 Town,C.D., Van Aken,S., Uteback,T., Koo,H. and Fraser,C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 UNPUBLISHED
 OTHER_GSS: BONHW74TF
 CONTACT: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..809
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="T01000DH3"
 /db_xref="taxon:3712"
 /clone="BONHW74"
 /clone_1lb="BO_1.6_2_KB_tot"
 /note="Vector: pHO51; Site 1: BstXI; 1.6-2 kb sheared
 total DNA inserted into pHO51 using BstXI linkers"

BASE COUNT 241 a 216 c 146 g 206 t
 ORIGIN
 Query Match 25.0%; Score 502.6; DB 29; Length 809;
 Best Local Similarity 86.8%; Pred.No. 6.3e-90;
 Matches 553; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 1342 TTGAGTATTTATTCACAGAGATGATATTCAGAGAACAGACAGTGGCGTCTATCG 1401
 DB 758 TTATATTTATGTCACAGAGATGATATTCAGAGAACAGACAGTGGCGTCTATCG 699
 QY 1402 AGTGAATCTCGCTAGATGCTCTTCAATCCAGATATGCAATCAACGATCAAAAACGAGC 1461
 DB 698 AGTGAATCTCGCTAGATGCTCTTCAATCCAGATATGCAATCAACGATCAAAAACGAGC 639
 QY 1462 TGAATCAAGTATGTCGGAATCAAGAGCCCTAGATGATGATCTGACTTGGCTTCACTTCAT 1521
 DB 638 TTGATTTGGTGTGGAAATCAAGAGCTGTAGATGATGATCTGACTTGGCTTCACTTCAT 579

QY 1522 ATCTAACGGCTGTGTGTAAGAGATATTAGGCTTCACTCTCCAGGCCCCACTTCTATCAT 1581
 DB 578 ATCTAACGGCTGTGTGTAAGAGATATTAGGCTTCACTCTCCAGGCCCCACTTCTATCAT 519
 QY 1582 GGGCCCGCTTGGCCATTAACAGACAGATGCTTGAATGCTGCTTGTCCGACAGGACCA 1641
 DB 518 GGGCCCGCTTGGCCATTAACAGATGCTTGAATGCTGCTTGTCCGACAGGACCA 459
 QY 1642 CAGCAATGTGAACATGTGGCCGCTATGCTATGCAACAGCTGTGGCTTATCTTTGG 1701
 DB 458 CAGCAATGTGAACATGTGGCCGCTATGCTATGCAACAGCTGTGGCTTATCTTTGG 399
 QY 1702 AGTTAAACCTGAGAGGTTGCTGGCAAAAGAGGTGAGTGGAGTTTCTGAGT 1761
 DB 398 AGTTAAACCTGAGAGGTTGCTGGCAAAAGAGGTGAGTGGAGTTTCTGAGT 339
 QY 1762 CGAATTTGAGACTTGACCTTTTCGGGTCGGTCTGAGATTGCCCCGGAGAAATCTTG 1821
 DB 338 CGAATTTGAGACTTGACCTTTTCGGGTCGGTCTGAGATTGCCCCGGAGAAATCTTG 279
 QY 1822 GTTTACTACCGCTTATGTTTGAACGGGATGATGTTAATGATTTGAATGGGACCGT 1881
 DB 278 GATTGACAAACGCTTACGTTGACCGGACGCTGTTACAGAAATTTGAATGGGACCGT 219
 QY 1882 CGAATGTGAACGGGCTTACTTATCTGAGAAATGAGGCTTTCTTGGCAGATGCTAATC 1941
 DB 218 CAGATGTGAACGGGCTTACTTATCTGAGAAATGAGGCTTTCTTGGCAGATGCTAATC 159
 QY 1942 CTCTTCTGCTTAATTTGGCCGCTAGGCGCACTTAAA 1978
 DB 158 CTCTTCTGCTTAATTTGGCCGCTAGGCGCACTTAAA 122

RESULT 4
 BH451907/c 674 bp DNA linear GSS 12-DEC-2001
 LOCUS BH451907 BOGS044TR BOGS Brassica oleracea genomic clone BOGS044, genomic
 DEFINITION survey sequence.
 ACCESSION BH451907
 VERSION BH451907.1 GI:17637618
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 674)
 Town,C.D., Van Aken,S., Uteback,T., Koo,H. and Fraser,C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 UNPUBLISHED
 OTHER_GSS: BOGS044TF
 CONTACT: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..674
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="T01000DH3"
 /db_xref="taxon:3712"
 /clone="BOGS044"
 /clone_1lb="BOGS"
 /note="Vector: pHO51; Site 1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pHO51 using BstXI linkers"

BASE COUNT 191 a 194 c 130 g 159 t
 ORIGIN

Query Match 23.8%; Score 479; DB 28; Length 674;
 Best Local Similarity 89.6%; Pred. No. 3.2e-85;
 Matches 515; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

1405 GGATCTCGTAGAGATGCTTCCTCATCATATGACATCAACAGGTAACAAACAGAGCTGG 1464
 674 GGATCTCGTAGAGATGCTTCCTCATCATATGACATCAACAGGTAACAAACAGAGCTGG 615
 QY 1465 ATCAAGTGTGGGAAATCAAGAGCCCTAAGTGAATCTGACTTGGCTTCACTTCATATC 1524
 Db 614 ATTCGATGTCGGAATAACAGAGCCCTGATGATGATGATGATGATGATGATGATGATG 555
 QY 1525 TAACGGCTGTGATGAAAGATATGAGGCTTCATCCCTCCAGGCCCACTTCATATGAGG 1584
 Db 554 TGACGGCTATGATGAAAGATGATGAGATGATGATGATGATGATGATGATGATGATG 495
 QY 1585 CCCGTTGGCCATTAACAGACGATCGTGTATGCTGCTTCTTCCGGCAGGACCAAG 1644
 Db 494 CCCGTTGGCCATTAACAGACGATCGTGTATGCTGCTTCTTCCGGCAGGACCAAG 435
 QY 1645 CAATGTGATGATGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1704
 Db 434 CAATGTGATGATGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 375
 QY 1705 TTAACCTGAGAGGTTCTGTCGAAAGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 1764
 Db 374 TTAACCTGAGAGGTTCTGTCGAAAGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 315
 QY 1765 ATTTAGAGCTTGACCTTCCGATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTG 1824
 Db 314 ACTTAGAGCTTGACCTTCCGATGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG 255
 QY 1825 TTACTAGCTTATGTTTGAAGGCGATGATGATGATGATGATGATGATGATGATGATGAT 1884
 Db 254 TGACTAGCTTATGTTTGAAGGCGATGATGATGATGATGATGATGATGATGATGATGAT 195
 QY 1885 ATGTAGAGGCTTACTTATGAGAACTGAGGCTTCTTGGAGATGAGCTATATCTC 1944
 Db 194 TTGGAAGCGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 135
 QY 1945 TTCTGCTAATTTGCGCGCTAGCGCAGTTAAAA 1979
 Db 134 TTGCTGCTAATTTGCGCGCTAGCGCAGTTAAAA 100

RESULT 5
 BH431683/c 804 bp DNA linear GSS 12-DEC-2001
 LOCUS BOHRT46R BOHR Brassica oleracea genomic clone BOHRT46, genomic
 DEFINITION survey sequence.

ACCESSION BH431683
 VERSION BH431683.1 GI:117617404
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea

REFERENCE
 AUTHORS Town, C.D., Van Aken, S., Uterback, T., Koo, H. and Fraser, C.M.
 TITLE Whole genome shotgun sequencing of Brassica oleracea
 JOURNAL Unpublished
 COMMENT Contact: Chris Town

TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers

FEATURES
 source 1..804

/organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="T01000D3"
 /db_xref="taxon:3712"
 /clone="BOHRT46"
 /clone_id="BOHR"
 /note="Vector: PHOS1, Site 1: BstXI, 2-3 kb sheared
 genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 241 a 182 c 155 g 226 t
 ORIGIN

Query Match 23.5%; Score 472; DB 28; Length 804;
 Best Local Similarity 77.1%; Pred. No. 8.1e-84;
 Matches 628; Conservative 0; Mismatches 150; Indels 36; Gaps 3;

761 AGTAGTAAGAACCCCTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 820
 778 AGTAGTAAGAACCCCTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 719
 QY 821 ATGTGCTGTATTTGGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 880
 Db 718 ATGTGCTGTATTTGGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 881 ATGTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 940
 Db 659 TTGTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 941 CTATCGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1000
 Db 599 TTATCGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 QY 1001 AACCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1060
 Db 539 AACCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 1061 GATTGCTGAGGTTTGTGCTTCCCTCATGATGATGATGATGATGATGATGATGATGATGAT 1120
 Db 479 GATTGCTGAGGTTTGTGCTTCCCTCATGATGATGATGATGATGATGATGATGATGATGAT 420
 QY 1121 ATGCGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1180
 Db 419 ATGCGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 385
 QY 1181 TGAAGATGCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1240
 Db 384 TGAAGATGCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 325
 QY 1241 TGTGAAATTAATAACATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1300
 Db 324 GATGGAATTAATAACATTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 265
 QY 1301 TTTAATACAGTTTCTAATAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1360
 Db 264 TTTAATACAGTTTCTAATAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 215
 QY 1361 GAT 1420
 Db 214 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 155
 QY 1421 GATCCTTCATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1480
 Db 154 GATCCTTCATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 95
 QY 1481 ATCAAGAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1540
 Db 94 ATCAAGAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 35
 QY 1541 AGAAGTATGAGGCTTATCCTCAGGCGCACTT 1574
 Db 34 AGAAGTATGAGGCTTATCCTCAGGCGCACTT 1

RESULT 6

Db 583 GATGTCCTTCATCAAGATACAGACACAGTTCAACAGAGCTTGACAGGTCGTAG 524
 QY 1478 GAATCAAGAGCCCTAGATGATATGACTTGGCTTCACTTCAATCAACGGCTGTGCT 1537
 Db 523 GAGATCAAGAGCCCTAGATGATATGACTTGGCTTCACTTCAATCAACGGCTGTGCT 464
 QY 1538 GAAAGAGTATGAGGCTTCACTTCCAGGCCCACTTCAATCAATGAGCCCGTTGGCCAT 1597
 Db 463 CAAAGAGTCTTGAGGCTTCACTTCCAGGCCCACTTCAATCAATGAGCCCGTTGGCCAT 404
 QY 1598 AACAGACAGATCGTTAGTGTGCTTGTTCGCGCAGGACCAACAGCAATGATGAACAT 1657
 Db 403 CACAGATACGATCAATTGACGCGCGTGTGTGCGCGAGGACCACTGCAATGATTAACAT 344
 QY 1658 GTGGCCGATCGATGATGATTCACACAGTGTGGGTTGATCTTGGAGTTTAAACCTGAGAG 1717
 Db 343 GTGGGCTATAGCGATGATTCACACAGTGTGGGTTGATCTTGGAGTTTAAACCTGAGAG 284
 QY 1718 GTTGTGCAAAAGAGTGTGAGTGTGAGTTTGTGCTTGGGTTGAGACTTGC 1777
 Db 283 GTTGTGAGCCAAAGAGTGTGAGTGTGAGTTTGTGCTTGGGTTGAGACTTGC 224
 QY 1778 ACCCTTGGGCTGTGCTGTGCTGATTTGCGCGGAGAAATCTTGTATACCTTAT 1837
 Db 223 ACCGTTGGGCTGTGCTGTGCTGATTTGCTGCGGAGAAATCTGAGTTGACACCGTGAC 164
 QY 1838 GTTGTGAGCGCGATGATGATTAAGTTTGAATG-----GACCGTCCAGTGTAA 1891
 Db 163 GTTGTGATCGGAGCGCTTGTGATGATTTGAATGCTTGCACCGTCCAGTGTAA 104
 QY 1892 CGGCTTACCTTATCTGGAACCTGAGGCTTCTTGTGAGATGCTATCCTCTGTC 1951
 Db 103 GACCGTTACCTTGTGCGGAAACGTGAGGCTCTGCTGTGAGATGCTATCCTCTGTC 44
 QY 1952 TAAATTGCGCGGTAGCGCAGTTAA 1977
 Db 43 TAAATTGCTGCTGCAAGCGCAGTATA 18

 RESULT 8
 BZ46758/c 659 bp DNA linear GSS 13-DEC-2002
 LOCUS BONND1LR BO.1.6.2 KB tot Brassica oleracea genomic clone BONND11,
 DEFINITION genomic survey sequence.
 ACCESSION BZ46758
 VERSION BZ46758.1 GI:26759933
 KEYWORDS
 SOURCE
 ORGANISM
 Brassica oleracea
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 659)
 REFERENCE
 AUTHORS Town,C.D., Van Aken,S., Uterback,T., Koo,H. and Frazer,C.M.
 TITLE Whole genome shotgun sequencing of Brassica oleracea
 JOURNAL Unpublished
 COMMENT Other GSSs: BONND1LR
 CONTACT: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 FEATURES
 Source
 1..659
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BONND11"

QY 78 TCATCTTAAGGCCCAAGGCC-AAAAGACACTCTCTTTTGT-----TATCTCTC 130
 Db 659 TCATCTTAAGGCCCAAGGCCCAAGAACACTCTCTTTTGTCTCTCTCTCT 600
 ORIGIN
 BASE COUNT 172 a 143 c 195 g 149 t
 Query Match 20.9%; Score 419.6; DB:29; Length 659;
 Best Local Similarity 82.2%; Pred.No. 2.3e-73;
 Matches 549; Conservative 0; Mismatches 99; Indels 20; Gaps 5;
 /clone_1id="BO.1.6.2_KB_tot"
 /note="Vector: PHOS1; Site: 1; BstXI, 1.6-2 kb sheared
 total DNA inserted into PHOS1 using BstXI linkers"
 QY 131 TTCTCTCACTTACTTCTCTTGAATCCGCCAT--GGCTACAAACTGAAAGC--TCCTT 186
 Db 599 ATAGCTTCACTTCACTTCTTGAATCCGCCATGAGGCTTACAGAACTGAAAGCTTTCCCT 540
 QY 187 AATCTTGGCCCTTTGTGCAATGAGGCTTCAAGCCAAACCAACTTGCCTTCCCT 246
 Db 539 GATCTTTGCCCTTCTCTCAATGAGGCTTCAAGCCAAACCAAGTCTGCTTTTCTAT 480
 QY 247 CCTGCGCGTCAACATCTGCTGCGCAATCTCTCTTTTATGAACTTATCCGGTGG 306
 Db 479 ACTTGCAATCTGATCTGCTGCGCTTCTCTCTCTTCTCTGAACTTATCCAGTGG 420
 QY 307 ACTGCTTGGGGGAAATACCTCTTGGCGGTTATATCCGTTTCAACAAACCGGAAA 366
 Db 419 ACTGATGAGGCAAAATCTCTTCTGCGCGTTAA-----CTAACAAACCGGAAAC 369
 QY 367 CGTATATCCGCTCCAAAGGCTTCCCTTGTGTTGAAAGCATGTCACTCATGTCAAGAC 426
 Db 368 GGTATATCCGCTCCAAAGGCTTCCCTTGTGTTGAAAGCATGTCTCATGTCCAGGCG 309
 QY 427 TCTAGTCAACCAAGATGCTGATGCAAGCTTGAATTCGAGCAAGAGGCTGATGGC 486
 Db 308 TCTAGCACACACCGATGCAAGAGCAACCGAGATGCAAGCCCAAGGCTGATGGC 249
 QY 487 TTTCAGCTTGAAGAGACTCGCGTATCTCTCACTGCAATCCCAAGTACGAAAGAT 546
 Db 248 TTTCAGCTTGAAGAGAGACAGCGTATCTCTCACTGCAATCCCAAGTACGAAAGAT 189
 QY 547 TCTGAATAGCCGGTTTTTGTCTGATGACCGGTTAAAGATGGGCTTCTCACTGATGT 606
 Db 188 TCTGAATAGCCGGTTTTTGTCTGACCGTCCGTTAAGATAGCTTACTCTCTGATGTT 129
 QY 607 TAAACAGCAATGATTTTGAACACACAGGTTTACTGGCAACGCTTGCCTGATGCG 666
 Db 128 TAAACAGCAATGATTTTGAACACACAGGTTTACTGGCAACGCTTGCCTGATGCG 69
 QY 667 TTGAACCATCTCTTATGACAAACAAATCGAAGAGCCGAGACCCACAGAGCTGAT 726
 Db 68 TTGAACCATCTCTTATGACAAACAAATCGAAGAGCCGAGAGCCGAGAGGAGTAT 9
 QY 727 CTCAGGCC 734
 Db 8 CGCTAGCC 1

 RESULT 9
 BZ443616 776 bp DNA linear GSS 13-DEC-2002
 LOCUS BONFP28TR BO.1.6.2 KB tot Brassica oleracea genomic clone BONFP28,
 DEFINITION genomic survey sequence.
 ACCESSION BZ443616
 VERSION BZ443616.1 GI:26703871
 KEYWORDS
 SOURCE
 ORGANISM
 Brassica oleracea
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 776)

AUTHORS Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.
 TITLE Whole genome shotgun sequencing of *Brassica oleracea*
 JOURNAL Unpublished
 COMMENT Other_GSSs: BONFP28TF
 Contact: Chris Town

TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.

FEATURES

source 1.776
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="T0100DH3"
 /db_xref="taxon:3712"
 /clone="BONFP28"
 /clone_1b="BO.1.6.2 KB tot"
 /note="Vector: PHOS1, Site 1: BstXI; 1.6-2 kb sheared
 total DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 196 a 156 c 207 g 217 t

ORIGIN

Query Match 19.7%; Score 396.6; DB 29; Length 776;
 Best Local Similarity 74.2%; Pred. No. 8.9e-69;
 Matches 529; Conservative 0; Mismatches 179; Indels 5; Gaps 2;

1230 ATGTGACGTTGTGGAATTAATAAGCAATTGAAATTTTGTGACATCAATTTG 1289
 64 ACGTTCCTATTATTTATTTAATAAATAAACAACAACATATATATATTTTAA 123
 1290 AATTTTATATTTTATTAACAGTTCTATATAAAGTAAAGTAAATTTTGTAGTTA 1349
 124 ATATATCCATTTTAGTATATGACACAGTTATGTATTAACGTAATTCATTTATG 183
 1350 TTATTACAGAGATGATATTCAGAGACAGACAGTTCGATCTTAATCGATGATC 1409
 184 GTTTTACAGAAATGATATTTAGAGACAGTCAACGTCGTGATTTGATCGAGTGAT 243
 1410 CTGCTAGATGTCCTTCATCCAGATGCAATCAAGGTCAAAAGAGCTGATCAA 1469
 244 CTGCTAGATGTCCTTCATCCAGATGCAATCAAGGTCAAAAGAGCTGATCAA 303
 1470 GTAGTCGAGAAATCAAGAGCCCTAGATGATCTGATCTGATCTTCAATATCTAAG 1529
 304 GTGTAAGGAGATCAAGAGCCCTAGATGATCTGATCTTCAATATCTAAG 363
 1530 GCTGTGATGAAAGATATTGAGCTTCATCTCCAGCCCACTTTATCATAGGGCCGT 1589
 364 GCTGTGATGAAAGATATTGAGCTTCATCTCCAGCCCACTTTGTCGGGCCGT 423
 1590 TTGGCCATACAGACAGATCTGTGATGTCCTTGTCCGGCAGGACCAAGCAATG 1649
 424 TTGCAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 483
 1650 GTGAACATGTCGCGGATGATGATGATGATGATGATGATGATGATGATGATG 1709
 484 GTTAACATGTCGCGGATGATGATGATGATGATGATGATGATGATGATGATG 543
 1710 CCTGAGAGGTCGTCGCAAAAGAGGATGATGATGATGATGATGATGATGATG 1769
 544 CCGGAAGGTTTGTAGCAAGAGAGGATGATGATGATGATGATGATGATGATG 603
 1770 AGACTGACCTTTCCGGTCTGTCGATTTGCGGATTTGCGGCAAGATCTTGTTACT 1829
 604 AGCTTTCACCGTTTGGGTCGGGTCGTCGATTTGTCGGAAGAACATGAGGTTTGA 663
 1830 ACCGTTATGTTTGGACGGGATGATGATGATGATGATGATGATGATGATGATG 1886
 664 ACCGTGACGTTTGGATGCGGACGCTCTTGATGATGATGATGATGATGATGATG 723

1887 GGTAAAGG--CGTTGACTTATCTGAGAACTGAGGCTTTCTTGCAGATGCT 1937
 724 GATGACAGACCGTTGACTTCTCGAGAACTGAGGCTCTGTCGTGATGCT 776

RESULT 10
 BH451899 576 bp DNA linear GSS 12-DEC-2001
 LOCUS BOGSO44TR BOGS Brassica oleracea genomic clone BOGSO44, genomic
 DEFINITION survey sequence.
 ACCESSION BH451899 GI:17637610
 VERSION BH451899.1
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 576)
 Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.
 TITLE Whole genome shotgun sequencing of *Brassica oleracea*
 JOURNAL Unpublished
 COMMENT Other_GSSs: BOGSO44TR
 Contact: Chris Town

TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.

FEATURES

source 1.576
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="T0100DH3"
 /db_xref="taxon:3712"
 /clone="BOGSO44"
 /clone_1b="BOGS"
 /note="Vector: PHOS1, Site 1: BstXI; 2-3 kb sheared
 genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 137 a 162 c 114 g 163 t

Query Match 19.0%; Score 382.6; DB 28; Length 576;
 Best Local Similarity 83.1%; Pred. No. 5.5e-66;
 Matches 466; Conservative 0; Mismatches 79; Indels 16; Gaps 2;

29 TGTATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 88
 25 TTTATTTTAAAGATTCGATTTTCTATTAATAAGCAAGATATCTGATCTTTAAG 84
 89 CCCCAGAGCCCAAAAGACACTCTCTTTTGTATCTCTCTCTCTCTCTCTCTCT 141
 85 CCCCAGAGCCCAAAAGACACTCTCTTTAGTTCTCTCTCTCTCTCTCTCTCT 144
 142 TACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 201
 145 TGCATCTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 204
 202 GTCCAAATGACAGGCTTCAAGCCAAACCAACCTTCTCTCTCTCTCTCTCTCT 261
 205 GTCCAAATGACAGGCTTCAAGCCAAACCAACCTTCTCTCTCTCTCTCTCTCT 264
 262 CATCTGCTGCGCATATCTCTCTCTTAATGACCTATCCGGTGGAGACTGTTGGGGA 321
 265 GCTGTGCTGCGCTTTTCTCTCTCTTGTGACCTTAATGACCTGATGATGAGGCA 324
 322 ATACTCTTCCGCGGTTAATATCCGTTCAACAAACCGGAAACGTTATTCGGGTC 381
 325 ATACTCTTCCGCGGTTAATATCCGTTCAACAAACCGGAAACGTTATTCGGGTC 375

QY	1340	TTTGTGATTATTATACAGAGATGATTTTCAGAGAACAGACACAGTTGGCGTCTTAAAT	1398
Db	93	TTTTTGTGGGTTTTACAGAAATGATTTATTCAGAGAACTGACACGGTGGCGTCTTGAAT	152
QY	1400	CGATGTGATCCTCGTAGATGCTCTTCATCCAGATATGATGATCAACGGTACAAACGCA	1455
Db	153	CGAATGTGATTTCTTGCTATGATGCTCTTCATCCAGATATTCAGTCAATGGTTCACACGCA	212

```

FEATURES
    source
        location/Qualifiers
            1..732
                /organism="Brassicica oleracea"
                /mol_type="genomic DNA"
                /strain="TO100DH3"
                /db_xref="taxon:3712"
                /clone="BOHCZ39"
                /clone_1fb="BOHC"
                /note="Vector: pHOSt, site_1: BstXI, 2-3 kb sheared

```

BASE COUNT genomic DNA inserted into PHOS1 using BstXI linkers"
 ORIGIN 196 a 164 c 169 g 203 t

Query Match 16.5%; Score 330.8; DB 28; Length 732;
 Best Local Similarity 75.2%; Pred. No. 1.2e-55;
 Matches 445; Conservative 0; Mismatches 132; Indels 15; Gaps 2;

QY 398 GTTGAAGCATGCTCATGTCAGACCTCTAGCTCAGCAGAAATGCTGAGCT 457
 DB 141 GTCCGAAGCATGCTCTTCATGTCACAGCTTACCTCAGCTGCTAGCCGCGG 200
 QY 458 GAGAAATTCGAGCCAGAGAGCTCATGCTTTCAGCTTACAGAGAGCTGCTGATGCTC 517
 DB 201 GAGAAATTCGAGCCAGAGAGCTTAAATGCGTTAGCTGAGAGATCTCGCTGATGCTC 260
 QY 518 ACCTGCAATCCGACGAGAGAGAGATCTGAAATAGCCGCTTTCGCTGATGAGCG 577
 DB 261 ACCTGCAATCCGAGCTGAGCTTAAAGAGATCTGAAATAGCTGCTGATGAGCG 320
 QY 578 GTTAAAGATCGCTTACTCATGATGTTTAAAGAGAGATGTTTTCAGCAGCGT 637
 DB 321 GTTAAAGATCGAGCTTACTCATGATGTTTAAAGAGAGATGTTTTCAGCAGCGT 380
 QY 638 GTTAACTGCGAGAGCTTCCGCTTTCAGCAGATCTCTTTAGTAAAGAAATC 697
 DB 381 GTTAACTGCGAGAGCTTCCGCTTTCAGCAGATCTCTTTAGTAAAGAAATC 440
 QY 698 AGAAGAGCCGAGAGAGAGAGAGATCTCAAGCAGAGATGTTGAGTTCTTGAATA 757
 DB 441 AAGGCTCTGAAAGAGAGAGAGAGATCTCAAGCAGAGATGTTGAGTTCTTGAATA 500
 QY 758 CAGAGTAAAGCA-----ACCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 811
 DB 501 CAGAGTAAAGCA-----ACCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 560
 QY 812 AAGCAATGATGCTGCTGTTATGCTGCAAGAGATGATGCTGTTGTTGTTGTTGTTGTT 866
 DB 561 AATTAATGATGCTGCTGTTATGCTGCAAGAGATGATGCTGTTGTTGTTGTTGTTGTT 620
 QY 867 -----TTGAGTTAGTGAATGCTGCAAGAGATGATGCTGTTGTTGTTGTTGTTGTT 922
 DB 621 GTGATGAGCTTCCGCTTTCAGAGAGATGATGCTGTTGTTGTTGTTGTTGTTGTTGTT 680
 QY 923 ACTGATCACTTCTCTGCTATGAGAGTTGATCTCAAGAGCTCCGCTGTA 974
 DB 681 ACCGATCACTTCTCTGCTATGAGAGTTGATCTCAAGAGCTCCGCTGTA 732

RESULT 13
 CB344213 826 bp mRNA linear EST 10-Apr-2003
 LOCUS CB344213
 DEFINITION CA48EN001_TAF_G02 Cabernet Sauvignon leaf-CA48EN Vitis vinifera
 cDNA clone CA48EN001_TAF_G02 5', mRNA sequence.
 ACCESSION CB344213.2 GI:29784005
 VERSION CB344213
 KEYWORDS EST.
 SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; Vitaceae; Vitis
 1 (bases 1 to 826)
 Goes da Silva, F., Lim, H., Iandolo, A., Baek, D., Jones, K., Walker
 M.A. and Cook, D.R.
 Transcriptional responses of Vitis vinifera to infection by the
 bacterial pathogen Xylella fastidiosa
 Unpublished
 On Mar 14, 2003 this sequence version replaced gi:2895180.
 Contact: Douglas Cook, PhD
 CARS Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA

FEATURES
 source
 location/Qualifiers
 1..826
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultiivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CA48EN001_TAF_G02"
 /sex="Hermaphrodite"
 /dev_stage="Berry stage I"
 /lab_host="Vitis vinifera"
 /clone_id="Cabernet Sauvignon leaf-CA48EN"
 /note="Organ: Berry; Vector: pDNR; Site: 1: Self; Site: 2:
 Self; CA48EN is a cDNA library of Cabernet Sauvignon
 leaves. The leaves were collected on July 25, 2001, in
 Napa Valley, California, and represent leaves in
 mid-season development. These leaves were asymptomatic and
 verified to be non-infected with the bacterial pathogen,
 Xylella fastidiosa, based on a diagnostic assay using PCR
 and Xylella-specific primer pairs. cDNA made by oligo-dT
 priming and directionally cloned. 5' and 3' adaptors were
 used in cloning as follows:
 5'-AAGCAGTGTATCAACGAGAGTGGCCATTACGCGCGGG-3' and
 5'-ATTCTAGAGCCGAGCGCGCGGAGATC-3' (30' 3' library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."

BASE COUNT 179 a 218 c 224 g 203 t 2 others
 ORIGIN

Query Match 15.5%; Score 311.2; DB 14; Length 826;
 Best Local Similarity 68.5%; Pred. No. 9.7e-52;
 Matches 430; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 1351 TATTACAGAGATGATATTCAGAGAGAGAGAGAGATGCTGTTTAAATGAGTATC 1410
 DB 36 TCTTATGGAAGATGATATTCAGAGAGAGAGAGAGATGCTGTTTAAATGAGTATC 95
 QY 1411 TCGCTAGATGCTCTCTTATCCAGATATGCAATCAAGCTTCAATCAATCAATCA 1470
 DB 96 TCGCAAGATGATGATCTTCACTGCTGCTCAATCAAGCTTCAATCAATCAATCA 155
 QY 1471 TAGTGGGAAATCAAGAGCCCTAGATGATGATGATGATGATGATGATGATGATG 1530
 DB 156 TCGTGGGAAATCAAGAGCCCTAGATGATGATGATGATGATGATGATGATGATG 215
 QY 1531 CTGTGGTGAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1590
 DB 216 CGGTGCTGAAGAGGTTCTGAGGCTGACCCACCGGGCCCACTGCTGCTGAGGCT 275
 QY 1591 TGGCCATACAGACAGATGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1650
 DB 276 TATCAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 335
 QY 1651 TGAACATGTTGGCCGCTATGATGATGATGATGATGATGATGATGATGATGATG 1710
 DB 336 TTAACATGTTGGCCGCTATGATGATGATGATGATGATGATGATGATGATGATG 395
 QY 1711 CTGAGAGGTTGTTGAGCAAGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1770
 DB 396 CCGACAGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 455
 QY 1771 GACTTGACCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1830
 DB 456 GGTCTGCGCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 515
 QY 1831 CGCTATGTTTGAAGAGGAGATGATGATGATGATGATGATGATGATGATGATGAT 1890
 DB 516 CAGCACCTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 575
 QY 1891 ACGGCGTGAATATCTGAGAAATGAGGCTTCTTTCGAGATGAGCTAACTCTTCTG 1950

Db 576 ACCGGGATGATCTATCCGAGTGTAGCTATCTCGGAAATGACACCCCTGACTG 635
 QY 1951 CTAAATGGCGCCGTAGCGCAGTTAAA 1978
 Db 636 TGAAGGATGGCGCCAGCGAATCATGATA 663

RESULT 14

LOCUS CB433993
 DEFINITION CB433993 638 bp mRNA linear EST 10-APR-2003
 C448N0001_11BF_G02 Cabernet Sauvignon leaf C448N vltis vinifera
 CNA clone C448N0001_11BF_G02 5', mRNA sequence.

ACCESSION CB433993
 VERSION CB433993.1 GI:28964960
 KEYWORDS EST
 SOURCE vltis vinifera
 ORGANISM Vitis vinifera

REFERENCE

1. Goes da Silva, F., Lim, H., Tandolino, A., Baek, J., Jones, K., Walker, M.A. and Cook, D.R.
 Transcriptional responses of Vitis vinifera to infection by the bacterial pathogen Xylella fastidiosa

JOURNAL COMMENT

Unpublished
 Contact: Douglas Cook, PhD
 CAES Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcoc@ucdavis.edu
 Seq primer: ACCGTACCGACATATGCC.
 Location/Qualifiers

FEATURES

source 1..638
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="C448N0001_11BF_G02"
 /sex="Hermaphrodite"
 /dev_stage="Berry stage I"
 /lab_host="DMSalpa"
 /clone_lib="Cabernet Sauvignon leaf C448N"
 /note="Organ: Berry; Vector: pDR, Site_1: Sfil; Site_2: Sfil; C448N is a cDNA library of Cabernet Sauvignon leaves. The leaves were collected on July 25, 2001, in Napa Valley, California, and represent leaves in mid-season development. These leaves were asymptomatic and verified to be non-infected with the bacterial pathogen, Xylella fastidiosa, based on a diagnostic assay using PCR and Xylella-specific primer pairs. cDNA made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
 5'-AAGCAGTGTATCAACGAGTGTGCGCTTATTCAGCGATGCC 1410
 5'-ATTCTAGAGCCGAGCGCGCCACATG-DT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."
 BASE COUNT 133 a 172 c 189 g 144 t
 ORIGIN

Query Match 15.1%; Score 304; DB 14; Length 638;
 Best Local Similarity 69.2%; Pred. No. 2.6e-50;
 Matches 415; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY 1351 TATTACAGAGATGATATTCAGAGAACACAGTGTGCGCTTATTCAGCGATGCC 1410
 Db 36 TCTTATCGGAGATGATATTCGGGGAAGGATACGGTGGACGTTTGTATAGATGGATAC 95
 QY 1411 TCGCTAGATGATGCTTCTATCCAGATTCATCAACGTTCAAAACAGCTGGATCAAG 1470

Db 96 TCGAAGATGATCTACACCCGTAGATTCCAAATCAAGGATCCACAGATGATGAGT 155
 QY 1471 TAGTGGGAAATCAAGAGCCCTAGATGATATGATCTTGGCTTACTTCCATATTAACG 1530
 Db 156 TGGTGGAAAGTCAAGCCCGGTAAGAGATCCGATTCGGCGATGATTAACCTGCGCG 215
 QY 1531 CTGTGGTAAAGAAATATTTAGAGCTTCACTCCAGCCCACTTATCAATGAGCCCGTT 1590
 Db 216 CGGTGATGAGAGGTTCTGAGGCTGACACCCAGCGGCCACTCTCTCGTGGGCCGTT 275
 QY 1591 TGGCATTAACACACACGATTCGTGATGTCCTTTTCCGCGAGGAGACACACATG 1650
 Db 276 TATCAATCAAGATTCACCGGTGATGCTATACAGTGCCTGCGGGTCCACGCGATG 335
 QY 1651 TGAACATGAGGCGCGTATCGATATCCACACGATGAGGATTCCTTTGAGATTTAAAC 1710
 Db 336 TTAACATGAGGCGCATCAACGAGATCCACAGATGTGTGATCCATCGACTTCACGC 395
 QY 1711 CTGAGAGATTCGTGCGAAAGAGATGAGTGTGATTTGGTCTTGGGTGGATTTGA 1770
 Db 396 CCGACGATTCGTGACACCCCTCTGACGTGAAATTCCTGTGATTCGATCTCA 455
 QY 1771 GACTTCAACCTTTCGGGTGCGGTGCGTGCATTCACCGGAAAGATCTTTTACTA 1830
 Db 456 GCGTGGCGCATTCGGGTGCGGCAAGAAACCTGCCCGAAAGACCTTGGCTTGACA 515
 QY 1831 CGGTATGTTTGAACCGCGGATGATGTTTACATGATGAGGACCGTCCGATGTA 1890
 Db 516 CAGCACCTTTTGGTGGCGTGGCTTTGATGATGATGATGATGATGATGATGATGATG 575
 QY 1891 ACGGCGTTGACTTATCTGAGAAAGTGAAGCTTTTTCGAGATGATCTTCTTCTG 1950
 Db 576 ACCGATGATCTATCCGAGTGTAGCTATCTCGGAAATGACACCCCTGACTG 635

RESULT 15

LOCUS BH581933 765 bp DNA linear GSS 15-DEC-2001
 DEFINITION BOHRM09TR BOHR Brassica oleracea genomic clone BOHRM09, genomic survey sequence.

ACCESSION BH581933
 VERSION BH581933.1 GI:17834390
 KEYWORDS GSS
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea

REFERENCE 1. eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 1 (bases 1 to 765)
 Town, C.D., Van Aken, S., Utebback, T., Koo, H. and Fraser, C.M.
 Whole genome shotgun sequencing of Brassica oleracea

Unpublished
 Other_GSS: BOHRM09TR
 Contact: Chris Town
 TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.

FEATURES
 source 1..765
 Location/Qualifiers

/organism="Brassica oleracea"
 /mol_type="Genomic DNA"
 /strain="TO1000D3"
 /db_xref="taxon:3712"
 /clone="BOHRM09"
 /clone_lib="BOHR"
 /note="Vector: PHOS1, Site 1: BstXI, 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"
 BASE COUNT 197 a 167 c 192 g 209 t

Tue Jan 20 17:43:19 2004

us-10-022-025a-1.rst

Page 11

ORIGIN

Query Match 14.4%; Score 289.2; DB 28; Length 765;
Best Local Similarity 65.3%; Pred. No. 2.4e-47;
Matches 474; Conservative 0; Mismatches 243; Indels 9; Gaps 3;

```
QY 136 TTCCTTCTTCTTGGTTCATCGCATGGTACGAAAGCTCCTTAATCTTGC 195
Db 13 TTATTCAGATGTTTCTCTGGTATAGAACGAAATAGAAATCTGTGGCTTTGC 72
QY 196 CTTTTCACAAATGACAGCTTCTAAGCAAAACCAACCTTCTCTCTCCCGCT 255
Db 73 TCTTCATCAAAATGCAATATTTTCATTCAGAACATTTTCTCTTCATCGCGCT 132
QY 256 CACATTCATCTGCTCGCATATCTCTTCTTATGACCTATCCCGGTGACCTGCTG 315
Db 133 AGCCATCACTGGTTACCTTCACCATGTAATTTTGGTCTACCTGTGACCGGCTTG 192
QY 316 GGGAAATACCTCTTC---GGCCGGTAAATCCGGTTCATAC---AAACCGGAAACGT 369
Db 193 GGGAAATACCTTTTCACTCGCGGTTTACTTCGGGTGTAAACCAAAAAACAAGATCT 252
QY 370 TATTCCTGGTCCAAAGGCTTCCCTTGGTGGAGCATGTCACTCATGTCAAGCACTCT 429
Db 253 CATTCCTGGTCCGAGAGGGTTCCTCACTCGTGGAGCATGAGCTTAAGGTCAAGCCGCT 312
QY 430 AGCTACCGGACGATCGCTGATGACGTGAGAAATTCGAGCCAGAGGCTCATGGCTTT 489
Db 313 AGCTACCGGACGATGACGAGCTGCTGCGATGAGAAACGCAAGCGGCTCATGGGCTT 372
QY 490 CAGCTTAGAGAGACTCGCGTATGTCACGTGCATCCGACGTAGCGAAAGATTTCT 549
Db 373 TAGCTCGGTGATACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 432
QY 550 GAATAGCCGGTTTTTGTGATGACCGGTTAAAGATCGGTTACTCACTGATGTTAA 609
Db 433 GAACGCTCGGTGATGACGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 492
QY 610 CAGAGCAATGGTTTGGACCAACGCTGTTACTGGCGAACGCTTCCCGTATCGCTTC 669
Db 493 CCGAGCCATGGGATTTGCTCCAAGTGTGCTTACTGCGCAAGCTACGTCGCTTAGGTTTC 552
QY 670 GAACCATCTCTTTAGTACAAACAAATGAGAGCCGAGCAACGACGATGATCTC 729
Db 553 GAACCTATCTTTTAAACCAAGCAATCAAAGCTGAGAGCAAGACGATGATGCTC 612
QY 730 AAGCCAGATGTTGATGTTCTTGAATAACAGATGTAAGAACCTGTTTGTCTGTA 789
Db 613 GACTCAGATGTAATGTAATTTGACGTAAACGCTAAACCG---CGTTGAGATGCTGA 669
QY 790 GTTGTCTTAAACGGGCTCGTTTAAACAATGATGTGCTCTGATTTGCGACAAGAGTGA 849
Db 670 TTTCCTPAAAAACGGCGTGTGTTGCTAAATGATGGGTTAGTTTCGGGAAACAGATGA 729
QY 850 GCTTGA 855
Db 730 ACTGA 735
```

Search completed: January 16, 2004, 13:41:24
Job time : 4830 secs